THE
LONDON, EDINBURGH, AND DUBLIN
PHILOSOPHICAL MAGAZINE
AND
JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

NOVEMBER 1927.

LXXX. *Experiments on the Rate of Evaporation of Small Spheres as a Method of Determining Diffusion Coefficients. —The Diffusion Coefficient of Iodine.* By BRYAN TOPLEY *and* ROBERT WHYTLAW-GRAY *.

Introduction.

EVAPORATION into a still atmosphere from the surface of a sphere of volatile material is a process governed by diffusion. The object of this communication is to describe an experimental method whereby observations of the rate of evaporation may be utilized to determine the diffusion coefficient. The method applies to the case of liquids and solids whose saturated vapour pressure does not exceed about 1 mm. at the temperature of the experiment, and requires the vapour pressure to be known with accuracy.

Stefan's general theory of diffusion [1] applied to the evaporation of a sphere of finite size suspended freely at the centre of a spherical shell of absorbing material which maintains a zero concentration of the vapour at the surface of the absorbing material, leads to the result :

$$dm/dt = \frac{4\pi \mathrm{M} \cdot \mathrm{D} \cdot \mathrm{P}}{\left(\frac{1}{r} - \frac{1}{r_0}\right)\mathrm{RT}} \cdot \log_e \left\{ \frac{1}{1 - p/\mathrm{P}} \right\},$$

* Communicated by the Authors.

where dm/dt = the rate of evaporation in gms. per second.

M = the molecular weight of the vapour.

D = the diffusion coefficient of the vapour in gms. per second per sq. cm. at the total pressure P.

p = the saturation pressure of the vapour in dynes per sq. cm.

P = the total (constant) pressure of the atmosphere between the sphere and the absorbent in dynes per sq. cm.

r = the radius of the evaporating sphere.

r_0 = the radius of the spherical shell of absorbent.

R = the gas constant in ergs per degree.

T = the absolute temperature.

By differentiating, this can be re-written in terms of the surface of the sphere :

$$\int_{S_2}^{S_1}\left(1-\frac{\sqrt{S}}{2r_0\sqrt{\pi}}\right)dS = \frac{8\pi\,M\,.\,D\,.\,P}{\rho\,.\,RT}\cdot\left(\log_e\left\{\frac{1}{1-p/P}\right\}\right)\cdot(t_2-t_1),$$

where S = the surface area, in cm. sq., of the sphere at t secs.

ρ = the density at T° of the sphere.

Thus, on plotting $\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]$ against t, a straight line should be obtained, the slope of which determines the diffusion coefficient.

It was noticed originally by Sresnewsky [2] that the rate of evaporation of a spherical droplet in air is directly proportional to its radius, and not to its surface ; this was confirmed by Morse [3], using solid spheres of iodine resting on the pan of a microbalance. Langmuir, in a paper discussing Morse's experiments, argued that if the evaporation of a small object in still air is controlled by diffusion and is analogous to the loss of heat from bodies of the same shape, then the rate of loss of mass of a freely-suspended sphere of volatile material should be given by the formula :

$$-\frac{dm}{dt}=\frac{4\pi\,M\,.\,D}{RT}\cdot p\,.\,r.$$

The equation deduced from Stefan's theory reduces to this expression for small values of the ratio p/P, and when r becomes infinite. Using this formula, Langmuir [4] calculated the value of D for iodine diffusing in air at atmospheric pressure, and at 20° from Morse's experiments, obtaining as a rough approximation the value ·07 ; this, he pointed out, is of the same order as the diffusion coefficient of other substances of high molecular weight, and therefore the result lends support to the view that evaporation under these conditions is determined by diffusion only. The equation,

$$\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_{S_2}^{S_1}=\frac{8\pi M\,.\,D\,.\,P}{\rho RT}\cdot\left(\log_e\left\{\frac{1}{1-p/P}\right\}\right)\cdot(t_2-t_1),$$

applied to a pure substance evaporating symmetrically into a still atmosphere at constant temperature, relates the rate of evaporation to the diffusion coefficient at the pressure P and to the saturated vapour pressure of the pure substance in question ; hence the experimental measurement, under carefully controlled conditions, of the rate of evaporation should serve to determine the diffusion coefficient if the vapour pressure is known ; alternatively, the vapour pressure if an independent measure of the diffusion coefficient is available. A correction which takes account of the self-cooling of the evaporating sphere is discussed later.

When the radius of the sphere is small compared with the distance of the absorbent, the term $\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}$ can be neglected ; so that for small values of the ratio p/P the equation becomes

$$-\frac{dS}{dt}=\frac{8\pi M\,.\,D}{\rho\,.\,RT}\cdot p.$$

As a close approximation, therefore, dS/dt should remain constant as the sphere evaporates. That this is true in practice was shown by Langmuir in connexion with Morse's experiments already mentioned, and in addition it has been found to hold for droplets of various pure liquids (see Whytlaw-Gray and Whitaker [5]).

A complete test of the relationship involves verifying experimentally the magnitude of the constant $\frac{8\pi M}{\rho\,.\,RT}$; for this purpose measurements are required on a substance for which both the vapour pressure and the diffusion coefficient are known. The number of slightly volatile substances for

which sufficiently accurate data for the vapour pressure exist is very limited ; iodine, however, seemed to be a favourable case, since the vapour pressure is known with accuracy in the region of room temperature, and is such that the evaporation takes place at a rate convenient for measurement. Moreover, the chemical reactivity of iodine vapour makes it experimentally easy to measure the diffusion coefficient independently. For this reason it was decided to repeat Morse's experiments under strictly controlled conditions.

The experimental work consisted of two parts : first, the measurement of the rate of evaporation of small spheres of iodine at four temperatures ; and second, the independent determination of the diffusion coefficient at the same temperatures.

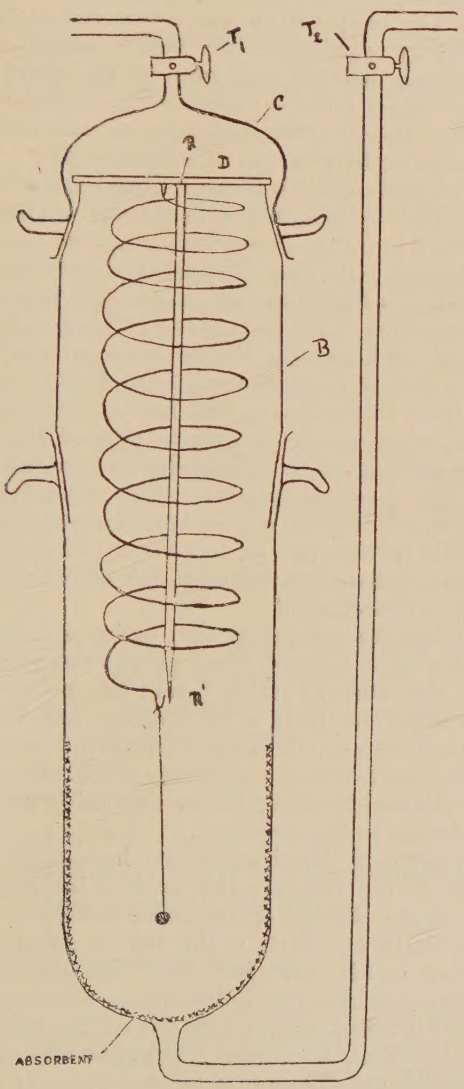
Measurement of the Rate of Evaporation of Iodine Spheres.

The diagram (fig. 1) shows the experimental arrangement used ; the essential part of the apparatus is a silica spiral spring made in the way described by H. Greville Smith [6]. The spring, together with a reference rod RR′, is sealed to the cross-bar D. The glass spring-balance case is made in three parts for ease of manipulation ; the cap C and the lower part of the cylindrical case are accurately ground on to the middle portion B. The particular spring used in the majority of the experiments was deliberately made comparatively insensitive, for two reasons : it was required to carry a relatively large load (up to 60 mgm.), and it was desired that the iodine sphere should evaporate completely without too great a change in the position of the sphere relative to the absorbent. A sensitivity of between 1 and 1·5 mgm. weight change for 1 mm. change in extension of the spring was found convenient ; the largest sphere investigated had an initial weight of 15 mgm., and the change in position when this evaporated completely was about 10 mm.

The spring was calibrated by the successive addition of aluminium wire weights, each of approximately 5 mgm. These were weighed separately on a Bender-Holbein microbalance, in terms of a set of N.P.L. standardized weights. In addition they were compared among themselves by observing the extension produced by hanging each weight in turn on a spiral spring ; this was repeated with a second spring. In this way the sensitivity of the spring actually used was found for the required range of extension relative to the reference point R′. Settings were made by a microscope supplied by the Cambridge Scientific Instrument Co.,

alternately on the reference point and on the hook which terminates the spiral spring.

Fig. 1.



The spring-balance case was carried by a stand designed to allow the balance to be lowered very smoothly under the

surface of the water of a large thermostat fitted with a plate-glass window. The stand, thermostat, and micrometer microscope were mounted on a heavy marble slab, which in turn was supported on thick rubber bungs, to lessen vibration. The temperature was controlled to $1/200^{c}$ by a sensitive Lowry regulator. The taps T_1 and T_2 were closed after the spring balance had been in the thermostat long enough for temperature equilibrium to be established; the pressure P of the air in the balance case was taken as equal to the barometric height read at that moment.

The Iodine Spheres.—Ordinary "Resublimed" iodine was further purified by grinding with KI and subliming off the iodine (twice), and then re-subliming a third time at as low a temperature as possible. The spheres were obtained in the following way: a fine silica fibre was dipped repeatedly in the just molten iodine, until from 50 to 100 mgm. had solidified on the end of the fibre in an irregular-shaped lump. This was fused into a drop of molten iodine by touching it momentarily with the tip of a fine flame; this has the effect of causing a sudden vaporization of the greater part of the iodine, the remainder melting in the atmosphere of iodine vapour to a more or less spherical droplet, which solidifies to a small sphere; it was found that a good approximation to a spherical droplet could be obtained up to a weight of about 25 mgm.; droplets larger than this tended to fall off while still molten, or to become pear-shaped.

There exists the possibility that iodine solidifies from the molten state in an unstable form which subsequently changes over into the ordinary stable form; the point was directly tested by experiments in a dilatometer, and some evidence was obtained for a small change in density in solidified iodine during the first few hours after its solidification. Sufficient time was allowed to elapse before beginning the evaporation for this change to be completed.

A thin layer of fused potassium hydroxide was used to absorb the iodine vapour; this adjusts the partial pressure of water-vapour in the spring-balance case to that of the lowest hydrate of KOH; the iodine spheres were thus evaporating in air with less than the normal amount of water-vapour, while the independent determinations of D carried out in tubes refer to the air of the laboratory, since metallic absorbents were used. This, however, could scarcely make a sufficient difference to be observed. In a few experiments copper gauze was used instead of potassium hydroxide; no difference could be traced to the change.

The correction for the self-cooling of the evaporating sphere

must now be considered. In the present experiments the vapour pressure of the iodine at the highest temperature (30°) is less than 0·5 mm., and the pressure of the air through which the iodine diffuses is atmospheric, so that $P \cdot \log_e \left\{ \frac{1}{1-p/P} \right\}$ may be taken as equal to p with sufficient accuracy. The diffusion coefficient is then calculated from

$$\frac{\left[S - \frac{\sqrt{S^3}}{3r_0\sqrt{\pi}} \right]_{S_2}^{S_1}}{(t_2 - t_1)} = \frac{8\pi M}{\rho \cdot RT} \cdot (D \cdot p).$$

The left-hand side of this equation is an experimentally measured quantity, which should be constant over the "life" of the evaporating sphere, provided that p remains constant; this depends upon whether the cooling of the sphere below the thermostat temperature is independent of the surface area of the sphere.

The extent of the self-cooling is readily calculated by equating the rate at which heat flows into the sphere to the rate at which heat is being rendered latent by the evaporation

$$dH/dt = -\frac{dS}{dt} \cdot \sqrt{S} \cdot \frac{\rho}{4\sqrt{\pi}} \cdot \frac{\lambda}{M},$$

where dH/dt = the rate at which the sphere is absorbing heat, in calories per second,

λ = the latent heat of evaporation, in calories per gm. molecule.

Heat can reach the sphere in three ways:

(1) By conduction through the gas;

(2) by conduction through the suspending fibre;

(3) by radiation from the absorbent.

The complete equation is then

$$-\frac{dS}{dt} \cdot \sqrt{S} \cdot \frac{\rho\lambda}{M \cdot 4\sqrt{\pi}} = [T_2 - T_1] \left\{ \frac{4\pi k}{\frac{2\sqrt{\pi}}{\sqrt{S}} - \frac{1}{r_0}} + \frac{\gamma\pi a^2}{r_0} \right\}$$

$$+ S \cdot 2\pi \int_0^\infty \{ f(\nu, T_2) - f(\nu, T_1) \} \alpha_\nu d\nu$$

where T_1 = the temperature of the sphere.

T_2 = the temperature of the absorbent.

k = the specific conductivity of the gas through which diffusion takes place.

a = the radius of the suspending fibre.

γ = the thermal conductivity of the material of the fibre.

α_ν = the absorption coefficient of the material of the sphere for radiation of frequency ν.

$f(\nu, T)$ = the Planck radiation function.

The term $\frac{\gamma\pi a^2}{r_0}.[T_2-T_1]$ representing the conduction through the fibre is not quite exact, since the temperature gradient in the gas surrounding the fibre is not linear. If the (silica) fibre is sufficiently fine, the conduction through the fibre may be neglected except when the sphere becomes very small, as is shown by the calculation given after the data of Exp. 32 (Table I.).

The term representing the effect of radiation assumes

TABLE I.

Exp. 32. Temp. 30·05° C. P = 761·1 mm.

S (cm.²).	$\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]$.	(t_2-t_1) seconds.	$\frac{\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_{S_2}^{S_1}}{(t_2-t_1)}$.
·05134	·05025	0	
·04843	·04743	1140	$2{\cdot}47\times10^{-6}$
·04289	·04204	3180	2·58 ,,
·03070	·03019	7920	2·53 ,,
·02520	·02482	10020	2·54 ,,
·01465	·01448	14040	2·55 ,,
·00830	·00823	16920	2·48 ,,

Average value = $2{\cdot}525\times10^{-6}$

Hence $(Dp) = 49{\cdot}26$

and $(T_2-T_1) = 0{\cdot}481°$, $\therefore$ Temp. of sphere = 29·57° C.

taking $k = 6{\cdot}12\times10^{-5}$ and $p =$ 0·451 mm.
for air at 30·05° C.

and $\lambda = 15100$ cals. $\rho = 4{\cdot}93$ } for iodine at 30·05° C.

$\therefore$ D corrected to P = 760 mm. = ·0820.

$dH/dt = 8{\cdot}498\times10^{-6}$ calorie per second,
for $S =$ ·00830 cm².

Maximum value of conduction through the silica fibre
$= 6{\cdot}798\times10^{-8}$ cal. per sec.,
taking $a =$ ·005 cm.
$\gamma =$ ·0036.

TABLE I. (*continued*).

Exp. 28. Temp. 20·00° C. $p = 757{\cdot}8$ mm.

S (cm.²).	$\left[S - \frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_1$.	$(t_2 - t_1)$ seconds.	$\frac{\left[S - \frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_2}{(t_2 - t_1)}$.
·09061	·08805	0	
·08841	·08594	1800	$1{\cdot}17 \times 10^{-6}$
·08600	·08363	3840	1·15 ,,
·08095	·07878	8100	1·14 ,,
·07411	·07209	13560	1·18 ,,
·07108	·06930	16350	1·15 ,,
·06455	·06301	21840	1·15 ,,
·05591	·05467	29169	1·14 ,,
·04698	·04602	36960	1·14 ,,
·03550	·03487	46410	1·15 ,,
·02836	·02791	52980	1·14 ,,
		Average value =	$1{\cdot}15 \times 10^{-6}$

Hence $(Dp) = 21{\cdot}66$ ∴ Temp. of sphere = 19·775° C.

and $(T_2 - T_1) = 0{\cdot}225°$, and $p = {\cdot}1961$ mm.

taking $k = 5{\cdot}969 \times 10^{-5}$ for air at 20·00° C.

and $\lambda = 15100$ cals. $\rho = 4{\cdot}93$ } for iodine at 20·00° C.

∴ D corrected to P = 760 mm. = ·0824.

$dH/dt = 1{\cdot}456 \times 10^{-5}$ calorie per second,

for $S = {\cdot}09061$ cm.²

Maximum amount of heat radiation to the sphere

$= 2{\cdot}98 \times 10^{-6}$ cal. per sec.

selective absorption by the sphere of certain frequencies ; it cannot be evaluated without a knowledge of the infra-red absorption of iodine. However, by taking $\alpha_\nu = \alpha = 1$, an upper limit is obtained for the rate at which heat can reach the sphere in the form of radiation for a given temperature difference ; since if the sphere behaves as a perfectly black body, the net rate of absorption of radiation is given by

$$S \,.\, \sigma \,.\, [T_2^4 - T_1^4],$$

where σ = Stefan's constant = $1{\cdot}38 \,.\, 10^{-12}$ cal.cm.$^{-2}$sec.$^{-1}$.

The greatest value that $(T_2 - T_1)$ can have in a given experiment must be less than that obtained by calculating it on the assumption that only conduction through the gas is operating to carry heat to the sphere. In this way an upper limit is found for the relative magnitude of the radiation term. The data of Exp. 28 (Table I.) show that

even when the sphere is comparatively large the radiation is not very important, and it diminishes with S. Since iodine is a bad conductor of electricity, it is probable that the absorption coefficient for infra-red radiation of the wave-lengths which are important at room temperature is much less than unity, and consequently the importance of the radiation term is still further diminished.

As applied to these experiments with iodine, the equation can now be simplified to

$$[T_2-T_1] = \left\{\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_{S_2}^{S_1}\Big/(t_2-t_1)\right\}\times\frac{\lambda\,.\,\rho}{8\pi M\,.\,k}=\frac{(D\,.\,p)\lambda.}{RT_2k}$$

It is concluded, therefore, that the lowering of temperature is independent of the size of the sphere over the range for which measurements are made. This conclusion is borne out by the results tabulated below for two typical experiments, in which the observed values of $\left\{\left[S-\frac{\sqrt{S^3}}{3r_0\sqrt{\pi}}\right]_{S_2}^{S_1}\Big/(t_2-t_1)\right\}$ are seen to be constant while the sphere evaporates down to about one-tenth of its initial mass. The correction for the self-cooling of the sphere is then obtained by first calculating the value of the product (pD) from the experimental line; the lowering in temperature (T_2-T_1) is found from this value of (pD). The value of p appropriate to the *corrected* temperature of the sphere is then used to calculate D.

This leaves out of account the small change in D itself resulting from the fact that the atmosphere through which diffusion takes place is at a temperature varying between T_2 and T_1; but the difference produced by this is outside the experimental accuracy of the method. The possible effect of convection set up by the cooled sphere is discussed later.

The fluctuations about a mean of the figures in the last column in Exps. 32 and 28 are caused mainly by errors in setting the microscope, produced by slight irregularities in the glass of the spring-balance case. In any particular experiment the mean is probably accurate to within 0·5 per cent.

Strictly, the sphere should be kept at the centre of a spherical spell of absorbent: actually, since in these experiments the iodine sphere moves upwards as the spring contracts during the evaporation by as much as 1 cm., the absorbent was placed on the inside of the cylindrical case, the absorbent layer being continued to a height well above the sphere. It was calculated for the particular apparatus

used that the cylindrical layer of absorbent with its hemispherical cap was equivalent, sufficiently nearly, to a sphere of radius 2 cm.

The vapour pressure of iodine at the different temperatures employed has been found by interpolation from the data of Baxter, Hickley, and Holmes [7].

The density of solid iodine has been taken as 4·93, as a mean value for the four temperatures (International Critical Tables). The values for the diffusion constant (fully corrected) are contained in Table II.

TABLE II.

Evaporation of Iodine Spheres.

Temperature 30·05°.	Temperature 25·05°.
D at 760 mm.	D at 760 mm.
·0840	·0815
·0851	·0817
·0854	·0822
·0851	·0815
·0832	·0819
·0820	·0834
·0847	·0829
·0873	·0840
	·0824
Average ... ·0846	Average ... ·0824

Temperature 20·00°.	Temperature 14·00°.
D at 760 mm.	D at 760 mm.
·0817	·0760
·0822	·0789
·0829	·0820
·0823	·0779
·0831	·0782
·0824	
Average ... ·0824	Average ... ·0786

The Diffusion Coefficient of Iodine.

The data in the literature for the diffusion coefficient of iodine vapour in air are scanty and somewhat discordant. Apart from the rough calculation made by Langmuir from Morse's experiments, the only diffusion measurements in air appear to be those of E. Mack [8], who found the value D = ·108 in air at 25° and 760 mm. It seemed difficult to reconcile this figure with the experiments of Mullaly and Jacques [9] on the rate of diffusion of iodine vapour in pure nitrogen, at pressures ranging from 9 to 19 mm. and at 20°.

Mack, making allowance for the difference between the cross-section of the average air molecule and the nitrogen molecule, calculated that the diffusion coefficient to be expected in the air at 25° and 760 mm. would be ·075. The discrepancy is unexpectedly large. The method adopted in Mack's experiments was to allow the iodine vapour to diffuse from the surface of solid iodine contained in a small glass cup ; the ground edge of the cup was placed in contact with the ground edge of a glass tube through which the diffusion took place to an absorbing medium (soda-lime or charcoal) at the other end of the tube. The weight of iodine evaporating in a known time was found from the loss in weight of the glass cup containing the iodine. An attempt was made to repeat these experiments, but it was found impossible to get concordant results ; an appreciable leakage of iodine vapour takes place at the junction of the edge of the cup and the edge of the diffusion tube, because of the large gradient in the partial pressure of iodine vapour across the junction. When the diffusion tube was surrounded by a closely-fitting sleeve of copper foil, the escaping iodine produced a deposit of cuprous iodide on the copper opposite to the junction of the iodine cup and the tube ; the formation of this protective layer and the consequent decrease in the gradient of iodine vapour concentration across the junction explains the observation that the apparent rate of diffusion diminished as the duration of the experiment was increased. Possibly the high value for D found by Mack was caused by a similar leakage.

The experimental arrangement was therefore modified as follows :—Iodine vapour from the surface of iodine fused into a glass cup diffuses downwards through a silica tube, being absorbed at the bottom by a suitable absorbent; the diffusion tube and iodine cap are enclosed in a copper tube fitted with a tap and kept in a thermostat. The weight of iodine diffusing in a known time is found by weighing the silica tube plus absorbent before and after the experiment. Since the absorbent has to be weighed, soda-lime is unsuitable; charcoal suffers from the disadvantage that it gives up adsorbed air as it adsorbs iodine; suitable absorbents were found to be silver powder (prepared electrolytically, and by reduction of silver nitrate solution with formaldehyde) and copper powder.

The diffusion coefficient is calculated from Stefan's formula :

$$D = \frac{m \cdot l}{q \cdot t} \cdot \frac{1}{c \log_e \left(\frac{c}{c - cs}\right)},$$

where m = the mass in gms. of iodine diffusing in t seconds

μl = the length of the diffusion tube in cms.

q = the cross-section of the tube in cm.^{2}

c = the total (constant) concentration of the gas in the tube, in gm. molecules per c.c.

c_S = the partial concentration of saturated iodine vapour in gm. molecules per c.c.

Two different silica tubes were used :

(1) $q = 2{\cdot}758$ cm.^{2} Length = 15·5 cm.

(2) $q = 2{\cdot}560$ cm.^{2} Length = 9·5 cm.

The actual value of l in any particular experiment was less than the length of the tube used by the depth of the layer of absorbent at the bottom of the tube, and was measured to the nearest 1/100th cm.

The extent of adsorption of iodine vapour on to the inside of the silica tube during the diffusion was tested by blank experiments in which the iodine cup was left in position on the tube for the usual duration of an experiment, but no absorbent was placed inside. The gain in weight amounted to about 1/10th mgm., while the weight of iodine diffusing in an experiment was usually 10 to 15 mgms. To eliminate this error arising from adsorption, diffusion was allowed to go on for some hours before weighing the tube plus absorbent for the first time, thus saturating the silica surface before the experiment.

The results are given below in Table III. The diffusion coefficients have been reduced to the values which they would have at 760 mm. of air. Except in the last two experiments, the diffusion was done at atmospheric pressure. The last two experiments, at 14° and 38 mm. air pressure, are in agreement with the mean of those done at atmospheric pressure and at the same temperature ; since the actual rate of evaporation of the iodine is 20 times as great at 38 mm. as at atmospheric pressure, and any surface cooling of the solid iodine would be greatly magnified, the concordance between the results at the two pressures shows that any such cooling is negligible.

The experiments themselves supply evidence that all the iodine is removed as soon as it diffuses on to the surface of the metal absorbent, so maintaining there a zero concentration of iodine vapour : first, the period of diffusion was lengthened in a few cases (so that the amount of iodine absorbed was increased) without producing a decrease in

the value obtained for D. Thus the progressive saturation of the active surface does not proceed beyond a safe limit. Second, the different forms of silver and copper absorbent lead to the same numerical result for D.

Table III.

Experiments in Tubes.

Temperature 30·05°.

Absorbent.	Tube.	D at 760 mm.
Electrolytic Silver.	(2)	·0850
,, ...	(2)	·0854
,, ...	(2)	·0867
,, ...	(2)	·0855
Reduced Silver	(2)	·0852
,, ,,	(2)	·0870
,, ,,	(2)	·0849
Copper Filings	(2)	·0853
Average		·0856

Temperature 25·05°.

Absorbent.	Tube.	D at 760 mm.
Reduced Silver	(1)	·0802
Electrolytic Silver.	(1)	·0797
,, ...	(1)	·0789
,, ...	(1)	·0786
,, ...	(2)	·0793
,, ...	(2)	·0830
,, ...	(2)	·0796
Reduced Silver	(2)	·0807
,, ,,	(2)	·0827
Average		·0804

Temperature 20·00°.

Absorbent.	Tube.	D at 760 mm.
Electrolytic Silver.	(2)	·0781
,, ...	(2)	·0801
,, ...	(2)	·0794
,, ...	(2)	·0787
Copper Filings	(2)	·0815
,, ,,	(2)	·0774
Average		·0792

Temperature 14·00°.

Absorbent.	Tube.	D at 760 mm.
Reduced Silver	(2)	·0763
,, ,,	(2)	·0735
,, ,,	(2)	·0755
,, ,,	(2)	.0724
,, ,,	(2)	·0750
,, ,,	(2)	·0745 (38 mm
,, ,,	(2)	·0761
Average		·0748

Table IV.

Temperature °C.	D from "tube" experiments at P=760 mm.	D from "sphere" experiments at P=760 mm.	Mean value of D at 760 mm.
30·05°	·0856	·0846	·0851
25·05°	·0804	·0824	·0814
20·00°	·0792	·0824	·0808
14·00°	·0748	·0786	·0767

Discusssion of Results, and Sources of Error.

The experiments with evaporating spheres and the experiments in tubes are compared in Table IV. At 30·05°, 25·05°, and 20·00° the discrepancy between the two values for D is not much greater than would be anticipated from the variations in the results of individual experiments recorded in Tables II. and III. The variation in the results for the iodine spheres is probably to be attributed to two main causes first: the difficulty of obtaining a perfect sphere, and second, the effect of vibration of the apparatus, which upsets the symmetry of the diffusion gradient. Both of these would tend to make the value of D too large. In the tube experiments the main source of error is inaccuracy in the weighings.

The figure ·0786, obtained from the evaporation of spheres at 14·00°, is very probably too high because the longer duration of the experiments at this temperature made it impossible to complete an experiment during a part of the day when the laboratory was comparatively free from vibration through local traffic.

An important point is that the discrepancy between the "sphere" experiments and the "tube" experiments does not increase at the higher temperatures, although the degree of self-cooling is greater; this must mean that the correction for self-cooling is adequate, and that any convection currents set up by the cooling of the sphere are not sufficient to affect seriously the results. The fact that (both for the "sphere" and the "tube" experiments) the decrease in D between 25·05° and 20·00° is too small may be due to an error in the relative values assumed for the vapour pressure of iodine at these temperatures: Baxter, Hickley, and Holmes do not give measurements for the vapour pressure between 15° and 30°, and this gap was covered by interpolation, after smoothing their experimental data by means of the Ramsay and Young formula against Knudsen's formula for the vapour pressure of mercury as a reference substance.

The results of these experiments are not sufficiently exact to give an accurate measure of the temperature coefficient of the rate of diffusion, but an approximate figure may be obtained; assuming that D varies as T^n, then from the mean values of D in the last column of Table IV., $n=2·0$.

It is concluded from these experiments that the method of allowing small spheres suspended from a silica microbalance to evaporate under definite conditions is capable of giving results for the diffusion coefficient accurate to about 2 or 3 per cent., and under favourable conditions (*e.g.* with

liquid drops, which are more accurately spherical, and with greater freedom from vibration) to within 1 per cent.

We have pleasure in expressing our thanks to Dr. F. R. Goss for the use of a Bender-Holbein microbalance, and to Messrs. Brunner and Mond for a grant which defrayed part of the cost of the apparatus.

References.

(1) J. Stefan, *Wien. Ber.* lxv. p. 323 (1872).
(2) Sresnewsky, *Beibl. Ann. Phys. Chemie*, vii. p. 888 (1883).
(3) Morse, Proc. Am. Acad. Arts Sci. xlv. p. 363 (1910).
(4) Langmuir, Phys. Rev. xii. p. 368 (1918).
(5) Whytlaw-Gray & Whitaker, Proc. Leeds Phil. Soc. i. p. 97 (1926).
(6) H. Greville Smith, 'Nature,' cxvi. p. 14 (1925).
(7) Baxter, Hickley, & Holmes, J. A. C. S. xxix. p. 127 (1907).
(8) E. Mack, J. A. C. S. xlvii. p. 2468 (1925).
(9) Mullaly & Jacques, Phil. Mag. lxviii. p. 1105 (1924).

Summary.

(1) An experimental arrangement is described for measuring the rate of evaporation of freely-suspended spheres of volatile material.

(2) The calculation of the diffusion coefficient of the vapour, and the correction for the self-cooling of the sphere, are discussed.

(3) The diffusion coefficient of iodine vapour in air has been measured over the temperature range from 14° to 30°.

The University, Leeds.

LXXXI. *The Warming of Walls.* *By* A. F. Dufton, *M.A.*, *D.I.C.* (*The Building Research Station*) *.

1. IT is well known that some rooms are not readily warmed. Complete analysis of the warming of a room is not easy. The influence, however, of the fabric of a wall is shown by the time taken to warm the surface when heat is supplied to it at a constant rate.

2. For simplicity a homogeneous wall (of thickness d, conductivity k and diffusivity h) initially at uniform temperature is considered. The rate of heating chosen is $2T\,k/d$, twice that necessary to maintain the desired temperature difference T between the faces of the wall, and the temperature of the second face is assumed to be constant.

* Communicated by the Author.

In time t the temperature of the wall at distance x from the heated surface increases by

$$T\left\{2-\frac{2x}{d}-\frac{16}{\pi^2}\sum_0^\infty (2n+1)^{-2}e^{-(2n+1)^2\pi^2ht/4d^2}\cos(2n+1)\pi x/2d\right\}.$$

The rise at the surface, the limit of this as x tends to zero, is

$$T\left\{2-\frac{16}{\pi^2}\sum_0^\infty (2n+1)^{-2}e^{-(2n+1)^2\pi^2ht/4d^2}\right\}$$

and attains the value T when

$$\frac{16}{\pi^2}\sum_0^\infty (2n+1)^{-2}\,e^{-(2n+1)^2\pi^2ht/4d^2}=1,$$

approximately after a time $d^2/5h$.

For a nine-inch brick wall the time is 6 hours: for a (one inch) wall of wood affording the same insulation 10 minutes.

3. The warming of a laminated wall depends largely upon the material at the heated surface. A room with masonry walls is much more readily warmed if lined with wood or other insulation of small thermal capacity.

LXXXII. *Bubbles and Drops and Stokes' Law.* *By* W. N. Bond, *M.A., D.Sc., F.Inst.P., Lecturer in Physics, University of Reading* *.

[Plate XIX.]

Summary.

This paper extends Stokes' calculations for the slow rectilinear motion of a solid sphere through viscous fluid, in the way he outlined, to the case where the sphere is composed of fluid. Experiments on the rate of rise of air bubbles in water-glass and in golden syrup, and on the velocity of the fluid in the neighbourhood of the bubbles in the former liquid, are in substantial agreement with the prediction that the bubble should rise one and a half times as fast as it would if it were solid. A smaller number of experiments on air bubbles and drops of syrup in castor oil also show reasonable agreement with theory. But those on water bubbles in castor oil indicate an effect due to surface contamination.

* Communicated by the Author.

Incidentally, the measured velocities of the fluid near steel spheres give a reasonably good check on Ladenburg's correction for the effect of the walls of the containing vessel.

Introduction.

STOKES'* mathematical investigation of the slow motion of a solid sphere through viscous liquid is well known. It has been the basis of a most useful method of determining viscosities †, as well as of a method of estimating the diameter of very small spheres ‡.

There is good evidence to show that, as Stokes assumed, there is no slip of the fluid at the surface of a solid sphere, except when surrounded by a gas at low pressure. When, however, the slow motion of a spherical gas bubble (or drop of liquid) through another fluid is concerned, there will probably be in general a finite tangential velocity of the fluid just inside and outside the spherical surface. The circulation of the fluid inside the sphere would therefore have to be taken into account in deducing the terminal velocity of a rising gas bubble or rising or falling liquid drop. Further, it is likely that the spherical surface of contact between the two fluids might itself offer a resistance to such tangential velocities (at any rate if the surface were contaminated). These two facts are specifically mentioned by Stokes §, though he does not complete the working-out of the theory.

This paper contains: firstly, a development of the theory given by Stokes; secondly, an account of experiments on the terminal velocity of rising air bubbles and falling drops of liquid; and thirdly, an account of measurements of the velocity of the surrounding liquid in the neighbourhood of rising air bubbles.

Theory.

When a solid sphere of radius a moves with a uniform velocity V_∞ through a large extent of fluid of density ρ and viscosity μ, then, provided we have ‖

$$\frac{V_\infty a \rho}{\mu} < 0{\cdot}6, \quad \ldots\ldots \quad (1)$$

* 'Mathematical and Physical Papers,' iii. pp. 55–59.

† Gibson & Jacobs, Journ. Chem. Soc. cxvii. p. 473 (1920).

‡ J. J. Thomson, Phil. Mag. (5) xlvi. p. 528 (1898). Millikan, 'The Electron,' p. 46 etc.

§ *Loc. cit.* pp. 61–62.

‖ H. D. Arnold. See Millikan, 'The Electron,' p. 96.

turbulence will not occur and the kinetic energy of the liquid may be neglected. The general nature of this result may be predicted by dimensional considerations; and a similar condition should also be required for the flow inside a drop or bubble to be purely viscous. These conditions are supposed fulfilled in all the following theory.

In place of considering a falling drop or rising bubble, we will suppose the drop or bubble at rest and the liquid moving. Taking polar coordinates, with origin at the centre of the sphere and axis towards the arriving fluid, we have for the surrounding fluid a stream function of the form *

$$\psi = (Ar^{-1} + Br + Cr^2 + Dr^4) \sin^2\theta. \quad . \quad . \quad (2)$$

The components of the fluid velocity resolved in an axial plane along and perpendicular to the radius are respectively

$$\left.\begin{aligned} R &= 2(Ar^{-3} + Br^{-1} + C + Dr^2)\cos\theta, \\ \Theta &= (Ar^{-3} - Br^{-1} - 2C - 4Dr^2)\sin\theta. \end{aligned}\right\} \quad . \quad . \quad . \quad (3)$$

The components in the directions r, θ of the pressure of the sphere on the fluid are respectively

$$\left.\begin{aligned} P_r &= (12Aa^{-4} + 6Ba^{-2} + 12Da)\,\mu\cos\theta, \\ T_\theta &= (6Aa^{-4} + 6Da)\,\mu\sin\theta, \end{aligned}\right\} \quad . \quad . \quad (4)$$

and the resultant force exerted by the outer fluid on the sphere on account of the motion is

$$-8\pi\mu B. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

A set of equations similar to (2), (3), and (4), but with A, B, μ, etc. replaced by dashed letters, applies to the fluid inside. The conditions at a distance require

$$D = 0, \quad C = -\tfrac{1}{2}V_\infty ;$$

and for the velocity not to be infinite at the origin we have

$$A' = 0, \quad B' = 0.$$

The other constants are determined by the conditions of no radial velocity at the surface, continuity of tangential velocity, and equality of the tangential component of pressure.

Thus, when $r = a$, we have

$$R = 0, \quad R' = 0, \quad \Theta = \Theta', \quad T_\theta = T_\theta'.$$

(The last of these conditions will not be true if the surface

* Stokes, *loc. cit.* pp. 55, 56, 59, and 31.

itself resists tangential stresses, the effect being as if the viscosity of the fluid inside had been increased. Also the normal component of pressure at the spherical surface is not the same in the two liquids, so the sphere will be very slightly distorted.)

Hence we obtain

$$A = -\frac{\mu'}{\mu+\mu'} \cdot \frac{a^3 V_\infty}{4}, \quad B = \frac{2\mu+3\mu'}{\mu+\mu'} \cdot \frac{aV_\infty}{4},$$

$$C' = \frac{\mu}{\mu+\mu'} \cdot \frac{V_\infty}{4}, \qquad D' = -\frac{\mu}{\mu+\mu'} \cdot \frac{V_\infty}{4a^2}.$$

It will be convenient to transform equations (3) and (5), and at the same time to return to the case where the sphere moves through liquid which is at rest at a distance.

We then obtain for the liquid outside, when the centre of the sphere is just passing the origin,

$$\left.\begin{aligned} R &= V_\infty \left\{ k\left(\frac{3}{2}\frac{a}{r} - \frac{1}{2}\frac{a^3}{r^3}\right) + (1-k)\frac{a^3}{r^3} \right\} \cos\theta, \\ \Theta &= V_\infty \left\{ k\left(-\frac{3}{4}\frac{a}{r} - \frac{1}{4}\frac{a^3}{r^3}\right) + (1-k)\left(\frac{1}{2}\frac{a^3}{r^3}\right) \right\} \sin\theta. \end{aligned}\right\} \quad (6)$$

The force on the sphere due to the motion becomes

$$6\pi k V_\infty a\mu, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and equating this to the gravitational force $\frac{4}{3}\pi a^3(\rho'-\rho)g$, the velocity of free fall becomes

$$V_\infty = \frac{1}{k}\left\{\frac{2}{9}\frac{(\rho'-\rho)ga^2}{\mu}\right\} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where ρ' and ρ are the densities of the fluid inside and outside.

The equations (6), (7), and (8) do not assume continuity of tangential force and velocity at the surface of the sphere. But by assuming these conditions, used above to deduce A and B, we obtain

$$k = \frac{2/3 + \mu'/\mu}{1+\mu'/\mu}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Equations (6) to (9) show that when μ' is large compared with μ, $k=1$ and the whole reduces to the case of a solid sphere, with no tangential velocity at the surface of the sphere. When, however, μ'/μ becomes small, we have $k=\frac{2}{3}$, and the sphere moves one and a half times as fast as a solid sphere of the same density. The terms in a^3/r^3 have now vanished, and there is no tangential component of pressure at the surface of the sphere.

Before describing the experiments, the effect of the walls of the vessel containing the fluid must be considered. When a solid sphere falls down the centre of liquid contained in a vertical cylinder, the mean velocity during the middle third of the path, V, may (according to Ladenburg *) be used to deduce V_∞ thus:

$$V_\infty = V\left(1+2{\cdot}4\frac{a}{R}\right)\left(1+3{\cdot}3\frac{a}{h}\right), \quad . \quad . \quad . \quad (10)$$

where R is the radius and h the total length of the cylinder of liquid. Since it is only the terms in a/r in equations (6) which represent appreciable velocities at a distance, these are the only terms that will contribute measurably to the wall correction. Hence for a drop or bubble we may use a modified correction:

$$V_\infty = V\left(1+2{\cdot}4k\frac{a}{R}\right)\left(1+3{\cdot}3\,k\frac{a}{h}\right). \quad . \quad . \quad (11)$$

Experiments I. *Terminal Velocities.*

In these experiments a vertical brass box of internal cross-section 4·75 cm. square was used, the total length of liquid column being generally about 13 cm. Two adjacent sides of the box had plate glass windows fitted, and at the centre of the base was a glass tube for the production of air bubbles of suitable size. The rate of rise of the air bubbles (or of the movement of liquid drops) was measured with a travelling microscope. The diameter of the bubbles was found both by a horizontal travel of the microscope and also by multiplying the velocity of the bubble by the time it took to move a distance equal to its own diameter. Various forms of illumination were used, but chiefly a white central background with the sides dark. The temperature of the liquid was read with a sensitive thermometer at one corner of the box; and the viscosity of the liquid was found by timing the rate of fall of ball bearings (of various sizes) before and after each measurement on a bubble or drop.

The substantial accuracy of the wall correction and the constancy of the viscosity at different rates of shear were thus verified. Any slight lack of homogeneity of the liquid was made to have little effect, by the bubble and steel sphere traversing the same part of the liquid. The condition of equation (1) was satisfied in all experiments for the surrounding liquid, and also in all cases $V_\infty a\rho'/\mu'$ was small compared with unity.

* Ladenburg, *Ann. der Physik*, (4) xxii. p. 287 (1907); xxiii. p. 447 (1907).

Most of the experiments were carried out using air bubbles rising in commercial "water-glass" (sodium silicate), in water-glass slightly diluted with water, and in commercial "golden syrup." The viscosities of these liquids were from about 2000 to 400 c.g.s.

Discarding some of the first less accurate observations (which fluctuated more, but had about the same mean value), the results of these tests were as follows:—

Values of $1/k$.

	Air in Water-glass.	Air in diluted Water-glass.	Air in Golden syrup.
	1·58	1·13	1·47
	1·41	1·46	1·43
	1·45	1·40	1·42
	1·49	1·31	1·39
	1·30	1·67	1·46
	1·50	1·56	1·41
	1·38	1·37	
Means ...	1·44	1·41	1·43

The general mean of all the measurements gives $1/k = 1{\cdot}43$. This does not differ from the theoretical value of 1·50 (μ'/μ very small) by more than can be accounted for by the fluctuation of the results. But these fluctuations, though partly due to experimental errors, may quite possibly have been contributed to by varying amounts of contamination of the surface of the bubble. (It may be remarked in passing that the above values of $1/k$ had been determined before the theoretical value of 1·50 had been deduced.)

Later, a smaller set of experiments was carried out using air, water, and golden syrup in castor oil. The very small air bubbles gave a value of approximately $1/k = 1{\cdot}5$. The other results were as follows:—

Values of $1/k$.

	Water in Castor oil.	Golden syrup in Castor oil.
	1·16	—
	1·12	—
	1·23	1·08
	1·13	0·99
	1·14	1·11
Means	1·16	1·06

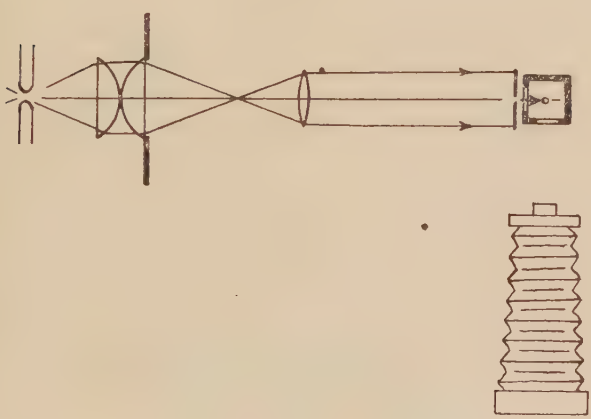
These would have been expected to be about 1·50 and

1·00 respectively. It must be presumed that surface contamination has prevented frictionless motion over the surface of the water-drop, and has on the other hand possibly allowed some slip at the surface of the drop of golden syrup.

Experiments II. *Motion of the Fluid near the Bubble.*

The vessel used in the previous experiments was now filled with "water-glass" in which a number of very small air bubbles had been produced. A horizontal parallel beam of light from an arc lamp fell normally on one of the two glass windows, passing through a narrow vertical slit fixed centrally on the window. In this way light fell only on all the small air bubbles in a vertical plane passing through the centre of the rising bubble experimented on. The motion of the bubble and of the very small bubbles in this plane was recorded photographically, using a camera facing the second window (shown diagrammatically in plan in fig. 1). Fig. 2

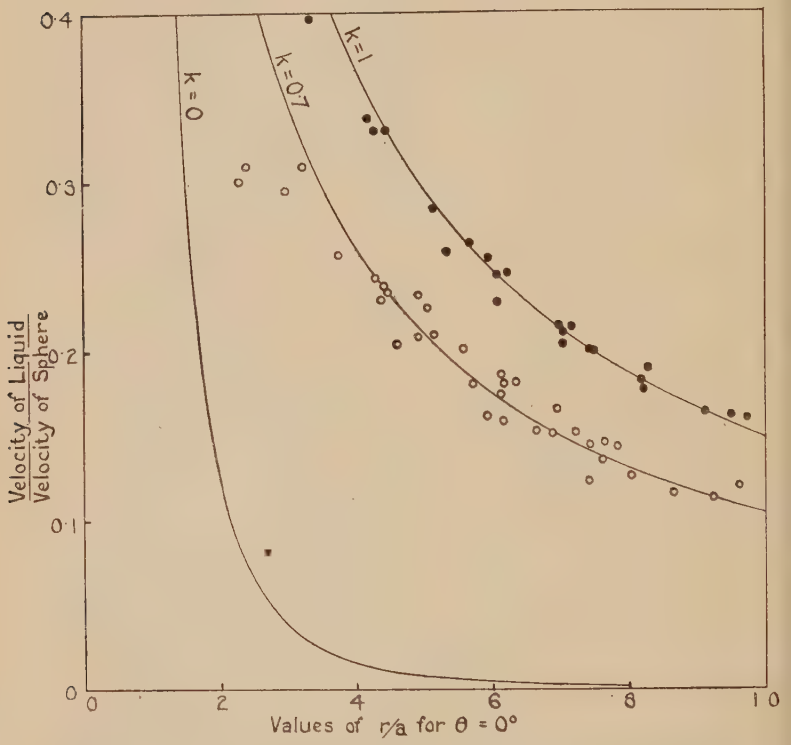
Fig. 1.



(Pl. XIX.) illustrates the type of result, successive exposures being made at equal intervals of time on the same plate. In this way exposures were made on three rising air bubbles and on a steel sphere slightly larger, and on one slightly smaller, than the bubbles. The diameters of the steel spheres were obtained previously, and also one of the air bubbles was measured by microscope. Using the measured camera

magnification, the diameters of image were deduced. As the form of illumination necessitated by this experiment was not suitable for obtaining a photographic record of the sizes of the spheres or bubbles themselves, a separate exposure was made near one end of four of the plates (fig 2, Pl. XIX.), showing the sphere or bubble in front of a white background. The diameters of the images thus obtained gave a satisfactory check with those predicted.

Fig. 3.



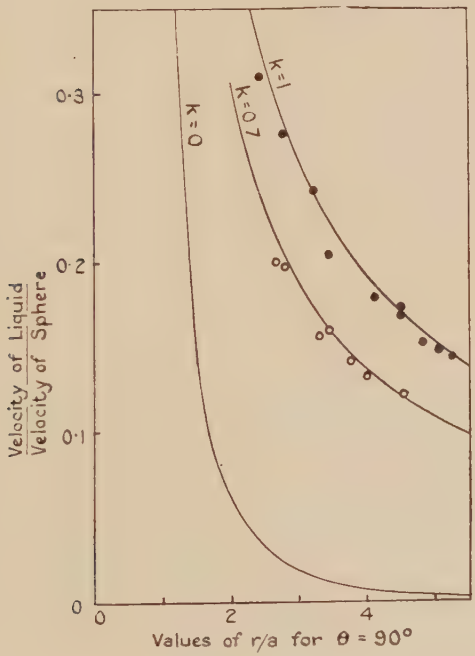
From the five photographs the velocity of the liquid at different distances from the spheres, at angles $\theta=0°$ and $\theta=90°$, was obtained as a multiple of the velocity of the sphere. On comparing the results obtained from the two steel spheres with those predicted by the usual theory (equations (6) with $1/k=1$), it was found necessary in each case to add a constant velocity to all the observations to obtain agreement. This velocity was slightly less than that

predicted by Ladenburg's results (equation 10), but showed evidence of increasing slightly near the sphere and probably reaching his value at the sphere.

The upper curves in figs. 3 and 4 are those predicted ($1/k=1$); and the results obtained from the steel spheres, after the addition of the above velocity on account of the wall correction, are indicated by dark dots.

The corresponding results for the air-bubble photographs were corrected according to the modified equation (11), assuming the value $1/k=1{\cdot}43$ previously obtained experi-

Fig. 4.



mentally (i. e., $k=0{\cdot}70$). These are represented by the circles in figs. 3 and 4. The curves passing through these circles were obtained from equations (6), again assuming $1/k=1{\cdot}43$.

The agreement between observation and the curves is quite good. (Even if the unmodified correction, equation (10), be used, all the air-bubble observations come

lower in these figures than the lowest steel observations at an equal value of r/a.)

The lowest curves in figs. 3 and 4 represent the velocity distribution for irrotational flow, obtained by putting $k=0$ in equations (6).

Thus the form of equations (6) has been verified, the value of k previously obtained found to be in agreement with the results, and incidentally a check has been obtained of Ladenburg's correction.

In conclusion the author would like to thank Prof. J. A. Crowther for the facilities that have made this work possible, and for the interest he has taken in the work; to Mr. J. S. Burgess, the laboratory steward, thanks are also due for his continued kind help in regard to apparatus.

Note added on October 8th, 1927.—From the above results it will be evident that when Stokes' Law is used to estimate the diameter of small drops of liquid, as in Millikan's determination of the electronic charge, e, a correction may have to be applied on account of the circulation of the liquid inside the drop. Thus, *unless skin friction decreases the effect*, we have to the required approximation

$$e\ (\text{corrected}) = e\left(1 - \frac{\mu}{2\mu'}\right).$$

Values of e obtained using mercury drops in air might therefore need reducing by a maximum of slightly more than $\frac{1}{2}$ per cent. I have not found data to deduce the maximum possible error in the case of Millikan's experiments using watch oil in air (though it seems likely that the maximum error might be only about $\frac{1}{20}$ per cent.).

This possible source of error does not seem to have been pointed out, and therefore it appears that it may not have been entirely eliminated.—W. N. B.

LXXXIII. *Ionization by Collision.* *By* L. G. H. Huxley, *B.A., Lecturer and Demonstrator, Electrical Laboratory, Oxford* *.

IN a previous number of this Journal (Phil. Mag. Sept. 1927, p. 505) J. Taylor has given a reply to my remarks on his "Photoelectric Theory of Sparking," in my communication entitled "Ionization by Collision" (Phil. Mag. May 1927, p. 1056), and, in addition, he also raises objections to the Theory of Sparking formulated by Townsend. It is proposed here both to consider these objections and to comment on the replies.

Taylor makes three distinctly different kinds of objection, which may be considered separately.

1. With regard to the theory of ionization by the collisions of electrons with molecules of the gas, Taylor raises an objection without giving reasons, stating that "the Townsend Theory of Ionization by collision for negative ions or electrons is not universally adopted"; and again, later, we find, with reference to the same hypothesis, "the photoelectric hypothesis was fitted into the existing fabric merely as an example of its application, but nowhere was a belief of the authenticity of Townsend's theory put forward."

It is to be presumed that Taylor, in proposing a theory of Sparking Potentials, assigns some action to the electrons; and it would be interesting to learn more precisely what is his theory of the action of the electrons and his reasons against the existing theory, which has so much experimental evidence to recommend it. To quote a definite experiment, it would be interesting to know how to explain the increase of conductivity obtained in pure helium by Bazzoni, as shown by Curve A, fig. 3 (Phil. Mag. Dec. 1916, p. 571).

2. Taylor's second objection is to the hypothesis that positive ions can ionize molecules of a gas in cases where the potential difference between the electrodes is of the order of the minimum sparking potential. He finds support for this objection in a paper by Sir J. J. Thomson (Phil. Mag. July 1924, p. 1), where reasons are proposed for believing that positive ions, unless they possess energies of several thousands of volts, do not ionize molecules of the gas. For instance, it is stated that to ionize a molecule of a gas having an ionizing potential of 10 volts, a hydrogen atom would require at least an energy corresponding to 4500 volts and an

* Communicated by Prof. J. S. Townsend, F.R.S.

oxygen atom 72,000 volts. In view of other phenomena, this certainly does not seem at all probable. For instance, the phenomena of thermo-luminescence in gases are adequately explained on the assumption that a gas molecule can produce the same effect on impact with another molecule as an electron having energy of the same order of magnitude. To quote Andrade ('The Structure of the Atom,' 3rd edit. p. 341) : "Now, it is reasonable to assume—by which we mean that the deductions from this assumption fit in well with observed facts, since there is no *a priori* inevitability about it—that the impact of another atom, provided it had the same energy, can produce the same effect as the impact of an electron."

Nor is it unreasonable to assume that a positively-charged molecule or atom can produce a similar effect. Since experiment shows that luminosity can be produced in the case of the alkali metals in the Bunsen flame, and that the energy to be communicated by the impacting body lies in the region of 3 volts, on Sir J. J. Thomson's hypothesis, a hydrogen atom in a Bunsen flame to produce this luminescence would require an energy in the neighbourhood of 2000 volts. The fraction of such atoms at 1700° C. whose translational energies exceed 2000 volts is of the order ϵ^{-6000}, a number which is so minute as to merit no further consideration. The presence of free electrons in flames has long been known *, and may obviously be explained by the same assumption. In fact, the generally-accepted theory of the luminescence and conductivity of hot gases is an extension of the hypothesis of the action of the positive ions to uncharged particles of the same order of mass.

It must again be emphasized that the main features of the electrical discharge in gases between concentric cylinders can most simply and accurately be explained on the hypothesis of ionization of gas molecules by collision with positive ions.

In my previous paper it was shown that the properties of the corona discharge indicate that, on Taylor's hypothesis, γ (the coefficient which occurs in Taylor's hypothesis) is constant, while the properties of discharges between parallel plates show a large variation of γ with the value of X (the electric force at the surface of the cathode). Taylor agrees with this conclusion, and shows that this result is still in accordance with his hypothesis ; for in the one case X is small and γ might quite well have fallen to a constant value,

* Harold A. Wilson, 'The Electrical Properties of Flames and of Incandescent Solids,' p. 89 (University of London Press).

while in the other case X is much larger and γ could in this region vary with X. It must be urged, however, that Townsend ('Electricity in Gases,' p. 368) was able to derive a formula connecting the critical force necessary to initiate a corona discharge in air with the diameter of the wire and the pressure, by using the numerical results yielded by experiments on sparking in uniform fields; so the theory indicates that the effects produced by electrons and positive ions are the same in both cases.

3. Taylor's third objection is to an "unmodified" Townsend Theory, which it is stated fails entirely to account for certain variations in sparking potentials due to electrode effects; but he omits to mention the fact that the electrode effects are actually dealt with in the treatise on Electricity in Gases. Townsend discusses two different effects of the positive ions—namely, the effect of positive ions in ionizing molecules of the gas, and their effect in setting free electrons from the negative electrode. He deals separately with these two possible effects of positive ions, and gives two formulæ, one involving only the effect of positive ions in ionizing molecules of the gas, and another involving only the effect of electronic emission from the cathode due to impinging positive ions. The following extracts from the treatise 'Electricity in Gases' serve to illustrate this point. On p. 330 we find a paragraph entitled "Effect of electrons set free from the negative electrode," in which the alternative theory of spark production is formulated. The two possible effects of positive ions are discussed, and it is stated at the end of paragraph 230 that "there are no reliable experiments to show the relative values of these two effects in a discharge for a given value of the product pS, but for simplicity it may be assumed that the gas effect predominates when pS exceeds the value corresponding to the minimum sparking potential." The effect of the alkali metals on the potentials in discharge-tubes is dealt with on p. 409, where the following statement is made: "The small cathode fall of potential is probably due to the fact that under the action of the radiation from the discharge or of the impacts of the positive ions, the alkali metals emit large numbers of electrons as compared with the numbers that would be emitted under similar circumstances from ordinary metals."

It is difficult to see why Taylor fails to find an explanation of electrode effects in the treatise on Electricity in Gases.

Dubois's results were quoted in contradiction to Taylor's; for with uncontaminated electrodes the one seemed to find

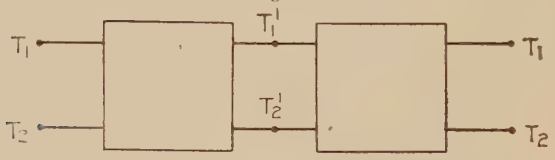
constancy in the sparking potential, while the other finds regular variations. It may be that Taylor's effect is more noticeable in the rare gases, with which Dubois does not deal. Having read Dubois's thesis, I was quite aware of his explanation of his results with saline contamination, and as far as the theory is concerned, he states that he reviews the existing theory, and gives one formula in which Townsend's two formulæ are combined.

Electrical Laboratory, Oxford.

LXXXIV. *An Extension of a Property of Artificial Lines.* *By* A. C. Bartlett, *B.A.** (Communication from the Research Staff of the General Electric Co., Wembley.)

IF any artificial line section having input and output terminals T_1 T_2, T_1 T_2 has within it a pair of terminals T_1' T_2', so that it may be represented, as in fig. 1, by two equal

Fig 1.



asymmetrical networks placed back to back, then, if the second half be removed, the impedance of the network remaining, measured at the terminals T_1T_2, is

$$Z_0 \tanh \frac{\theta}{2} \quad \text{if } T_1' \, T_2' \text{ be short-circuited,}$$

and $$Z_0 \coth \frac{\theta}{2} \quad \text{if } T_1' \, T_2' \text{ be open-circuited,}$$

where Z_0 and θ are the characteristic impedance and propagation constant of the artificial line section.

As an example, take the T section shown in fig. 2 (*a*). In fig. 2 (*b*) a half-section is shown, and it is seen that

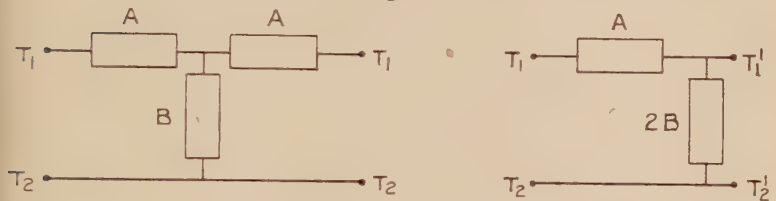
$$Z_0 \coth \frac{\theta}{2} = A + 2B$$

and $$Z_0 \tanh \frac{\theta}{2} = A.$$

* Communicated by C. C. Paterson.

It will be shown that this is a special case of a more general theorem which may be stated thus :—If an artificial line section has within it n terminals $T_1'\ T_2'\ \ldots\ T_n'$, such that, if these terminals are open, the artificial line section

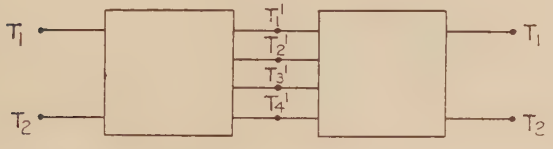
Fig. 2.



is cut into two exactly similar halves, the impedance of one-half measured at the terminals $T_1\ T_2$ with $T_1', T_2' \ldots T_n'$ free is $Z_0 \coth \frac{\theta}{2}$, and with $T_1', T_2' \ldots T_n'$ all connected together is $Z_0 \tanh \frac{\theta}{2}$.

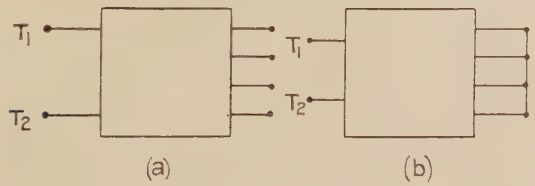
Thus, if fig. 3 represents the artificial line section (taking

Fig. 3.



$n=4$), then $Z_0 \coth \frac{\theta}{2}$ is the impedance of fig. 4 (*a*) and $Z_0 \tanh \frac{\theta}{2}$ of fig. 4 (*b*).

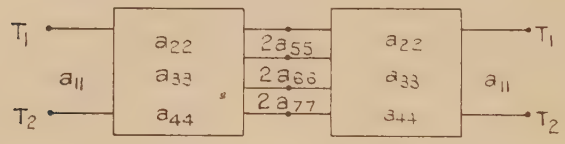
Fig. 4.



The proof, though simple, is rather long. To simplify

matters, suppose, as above, that $n=4$, and let the artificial line section be as shown diagrammatically in fig. 5.

Fig. 5.



Let a_{11} be the sum of the impedances round the terminal meshes ; let a_{22}, a_{33}, and a_{44} be the sums of impedances round meshes inside each half-section ; $2a_{55}$, $2a_{66}$, and $2a_{77}$ be the same for the meshes common to the two half-sections. In addition, there will be the mutual impedances a_{12} etc. between the meshes a_{11} and a_{22} etc. in each half.

Then the characteristic impedance will be given by *

$$Z_0^2=\Delta\Big/\frac{\partial^2\Delta}{\partial a_{11}\partial a_{11}},$$

where Δ is the axi-symmetric determinant

$$\begin{vmatrix}
a_{11}, & a_{12}, & a_{13}, & a_{14}, & a_{15}, & a_{16}, & a_{17}, & 0, & 0, & 0, & 0 \\
a_{21}, & a_{22}, & a_{23}, & a_{24}, & a_{25}, & a_{26}, & a_{27}, & 0, & 0, & 0, & 0 \\
a_{31}, & a_{32}, & a_{33}, & a_{34}, & a_{35}, & a_{36}, & a_{37}, & 0, & 0, & 0, & 0 \\
a_{41}, & a_{42}, & a_{43}, & a_{44}, & a_{45}, & a_{46}, & a_{47}, & 0, & 0, & 0, & 0 \\
a_{51}, & a_{52}, & a_{53}, & a_{54}, & 2a_{55}, & 2a_{56}, & 2a_{57}, & a_{54}, & a_{53}, & a_{52}, & a_{51} \\
a_{61}, & a_{62}, & a_{63}, & a_{64}, & 2a_{65}, & 2a_{66}, & 2a_{67}, & a_{64}, & a_{63}, & a_{62}, & a_{61} \\
a_{71}, & a_{72}, & a_{73}, & a_{74}, & 2a_{75}, & 2a_{76}, & 2a_{77}, & a_{74}, & a_{73}, & a_{72}, & a_{71} \\
0, & 0, & 0, & 0, & a_{45}, & a_{46}, & a_{47}, & a_{44}, & a_{43}, & a_{42}, & a_{41} \\
0, & 0, & 0, & 0, & a_{35}, & a_{36}, & a_{37}, & a_{34}, & a_{33}, & a_{32}, & a_{31} \\
0, & 0, & 0, & 0, & a_{25}, & a_{26}, & a_{27}, & a_{24}, & a_{23}, & a_{22}, & a_{21} \\
0, & 0, & 0, & 0, & a_{15}, & a_{16}, & a_{17}, & a_{14}, & a_{13}, & a_{12}, & a_{11}
\end{vmatrix},$$

in which $$a_{rs}=a_{sr}.$$

This determinant, which is partially centro-symmetric, can be simplified in the following way :—To the first row add the last row, to the second row add the last but one row, and similarly for the third and fourth rows. Next, in the resulting determinant :—From the last column subtract the first column, from the last but one column subtract the second column, and treat similarly the last but two and last but three columns. It is then found that the determinant

* *Cf.* J. I. E. E. pp. 223–227, Feb. 1927. Equation (20).

breaks up into the product of two other determinants giving

$$\Delta = 2^3 \Delta' \Delta'',$$

where $\Delta' =$

$$\begin{vmatrix} a_{11}, & a_{12}, & a_{13}, & a_{14}, & a_{15}, & a_{16}, & a_{17} \\ a_{21}, & a_{22}, & a_{23}, & a_{24}, & a_{25}, & a_{26}, & a_{27} \\ a_{31}, & a_{32}, & a_{33}, & a_{34}, & a_{35}, & a_{36}, & a_{37} \\ a_{41}, & a_{42}, & a_{43}, & a_{44}, & a_{45}, & a_{46}, & a_{47} \\ a_{51}, & a_{52}, & a_{53}, & a_{54}, & a_{55}, & a_{56}, & a_{57} \\ a_{61}, & a_{62}, & a_{63}, & a_{64}, & a_{65}, & a_{66}, & a_{67} \\ a_{71}, & a_{72}, & a_{73}, & a_{74}, & a_{75}, & a_{76}, & a_{77} \end{vmatrix}$$

and $\Delta'' =$

$$\begin{vmatrix} a_{11}, & a_{12}, & a_{13}, & a_{14} \\ a_{21}, & a_{22}, & a_{23}, & a_{24} \\ a_{31}, & a_{32}, & a_{33}, & a_{34} \\ a_{41}, & a_{42}, & a_{43}, & a_{44} \end{vmatrix},$$

Similarly,

$$\frac{\partial^2 \Delta}{\partial a_{11} \partial a_{11}} = 2^3 \frac{\partial \Delta'}{\partial a_{11}} \cdot \frac{\partial \Delta''}{\partial a_{11}},$$

so that

$$Z_0^2 = \frac{\Delta'}{\dfrac{\partial \Delta'}{\partial a_{11}}} \cdot \frac{\Delta''}{\dfrac{\partial \Delta''}{\partial a_{11}}}.$$

But $\Delta' \Big/ \dfrac{\partial \Delta'}{\partial a_{11}}$ is seen to be the impedance of the first half of the network of fig. 5 with all the terminals $T_1'\ T_2'\ T_3'\ T_4'$ connected together, while $\Delta'' \Big/ \dfrac{\partial \Delta''}{\partial a_{11}}$ is the impedance of the same network with $T_1'\ T_2'\ T_3'\ T_4'$ all free.

Since their product is equal to Z_0^2,

$$\Delta' \Big/ \frac{\partial \Delta'}{\partial a_{11}} \quad \text{and} \quad \Delta'' \Big/ \frac{\partial \Delta''}{\partial a_{11}}$$

may be written as

$$Z_0 \tanh \frac{\phi}{2} \quad \text{and} \quad Z_0 \coth \frac{\phi}{2}.$$

By using the identity

$$Z_0 \sinh \phi = 2 \Big/ \left(1/Z_0 \tanh \frac{\phi}{2} - 1/Z_0 \coth \frac{\phi}{2} \right),$$

and inserting values in the right-hand side, $Z_0 \sinh \phi$ can be found; if this is done, it will be found after reduction to be equal to Δ divided by the determinant obtained by omitting from Δ the first row and last column. But this is the receiving-end impedance of the original artificial line

section short-circuited at the receiving end, and is equal to $Z_0 \sinh \theta$; hence

$$\phi = \theta,$$

and the theorem is proved for this case.

The method, however, is quite general, and therefore the general theorem holds.

A number of simple corollaries follow for artificial line sections that can be bisected in the manner above :—

1. An exactly equivalent bridge section can be constructed.

2. By making bridge sections in which for one pair of arms any ladder network made up of any numerical multiples of $Z_0 \tanh \frac{\theta}{2}$ and $Z_0 \coth \frac{\theta}{2}$ is used, while for the other pair of arms the reciprocal of this network with respect to Z_0 is used, an indefinite number of bridge sections can be constructed having the same Z_0 as the original section but different propagation constants.

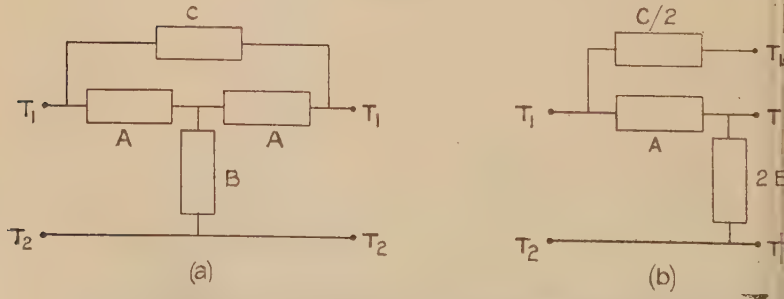
3. If the original section is modified by opening any of the terminals $T_1', T_2' \ldots T_n'$ and inserting any series impedances, then, if Z_0' and θ' are the constants of the modified section,

$$Z_0 \coth \frac{\theta}{2} = Z_0' \coth \frac{\theta'}{2}.$$

4. If the original section is modified by connecting any pairs of the terminals $T_1', T_2' \ldots T_n'$ by any impedances, then, if Z_0' and θ' are the constants of the modified section,

$$Z_0 \tanh \frac{\theta}{2} = Z_0' \tanh \frac{\theta'}{2}.$$

Fig. 6.



As a special case, consider the three-element artificial line section shown in fig. 6 (*a*). A half-section is shown in fig. 6 (*b*), three terminals, $T_1' T_2' T_3'$, being required for the bisection.

Thus for this artificial line section it can be seen from inspection that

$$Z_0 \coth \frac{\theta}{2} = A + 2B,$$

$$Z_0 \tanh \frac{\theta}{2} = \frac{1}{\frac{1}{A} + \frac{1}{C/2}} = \frac{AC}{2A+C};$$

and therefore

$$Z_0^2 = \frac{AC(A+2B)}{2A+C}.$$

This theorem for any artificial line to which it applies usually affords the simplest method of obtaining the constants in terms of the elements of the section.

From any artificial line section of the type considered, another can be derived by taking the two half-sections and connecting T_1' of one half-section to T_n' of the other half, T_2' to T'_{n-1}, etc. For this derived section the determinant Δ is completely centro-symmetric. The application, however, of the theorem that any centro-symmetric determinant is the product of two determinants does not appear to lead to any simple physical interpretation, though it does simplify the calculation of the constants.

LXXXV. *On Resonance in Pipes Stopped with Imperfect Reflectors.* *By* E. T. Paris, *D.Sc., F.Inst.P.**

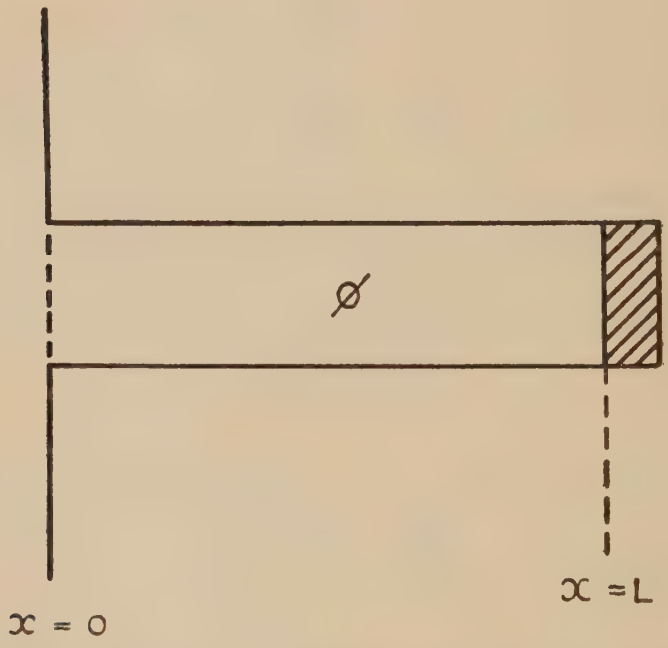
THE problem investigated in the following paper is that of resonance in a straight cylindrical pipe closed at one end by some material or mechanism which absorbs sound-energy. The pipe is supposed to have rigid walls and a diameter neither so great compared with the wave-length of the source that transverse vibrations are excited, nor so small that there is appreciable loss of energy from the operation of viscous forces. Thus the motion of the air within the pipe, due to a pure-tone source of sound, can be represented by two plane-waves of suitable amplitudes and phases travelling in opposite directions parallel to the axis of the pipe.

The first case to be considered will be that of a pipe under the influence of an external simple source of sound, the strength of which (measured in c.c. per sec.) varies with

* Communicated by the Author.

time in a simple harmonic manner. It will be supposed that the pipe is provided with an infinite plane reflecting flange lying in the plane $x=0$ (fig. 1), and that the source of sound is at such a distance from the pipe that the amplitude of the velocity-potential due to it is sensibly constant over the mouth of the pipe. This potential will be denoted by $2\mathrm{F}\epsilon^{\iota kat}$.

Fig. 1.



To allow for the escape of sound-energy from the open end, a velocity-potential may be assumed within the pipe of the form $\mathrm{A}(\beta\epsilon^{-\iota kx}+\epsilon^{\iota kx})\epsilon^{\iota kat}$, where β is the coefficient of reflexion at the open end. To this potential must be added a second potential, $2\mathrm{F}\cos kx\,.\,\epsilon^{\iota kat}$, to sustain the pressure-variations at the open end due to the external source of sound. Thus the total potential within the pipe is of the form

$$\phi=\{(\mathrm{F}+\beta\mathrm{A})\epsilon^{-\iota kx}+(\mathrm{F}+\mathrm{A})\epsilon^{\iota kx}\}\epsilon^{\iota kat}. \quad . \quad . \quad (1)$$

The value of A must be adjusted so as to make ϕ comply with the requirements of the problem in those parts of the pipe away from the open end.

It will now be supposed that the pipe is closed at $x=\mathrm{L}$ by some material or mechanism which absorbs sound-energy. In order to investigate the resulting velocity-potential within the pipe, use will be made of the Webster-Stewart conception of "acoustical impedance" *. According to Stewart the impedance of an acoustical receiver is the ratio $\delta p/(dq/dt)$, where δp is the pressure-variation at the receiver due to the source, and dq/dt is the rate of volume-displacement of air at the receiving surface. If ϕ is the velocity-potential (due to the source) at the receiver, $\delta p=\rho(d\phi/dt)$, so that acoustical impedance is the ratio $\rho\phi/q$ if both ϕ and q vary as $\epsilon^{i\omega t}$. To avoid an unnecessary repetition of ρ, impedance will here be taken to be the ratio ϕ/q, and its reciprocal, q/ϕ, will be denoted by Ω. It is convenient to have a name for Ω, and it will be referred to as "acoustical admittance."

Thus, if Ω is the admittance of the receiving-surface closing the pipe at $x=\mathrm{L}$, and dq/dt is the rate of volume-displacement of air across the plane $x=\mathrm{L}$, the condition to be satisfied by ϕ when $x=\mathrm{L}$ is

$$-\sigma\frac{\partial\phi}{\partial x}=\frac{dq}{dt}=\Omega\frac{\partial\phi}{\partial t}, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

σ being the cross-sectional area of the pipe. By combining (2) with (1), we find that

$$\mathrm{F}+\mathrm{A}=\mu\epsilon^{-2ik\mathrm{L}}(\mathrm{F}+\beta\mathrm{A}), \quad . \quad . \quad . \quad . \quad (3)$$

where

$$\mu=\frac{1-(a\Omega/\sigma)}{1+(a\Omega/\sigma)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Thus

$$\mathrm{A}=-\frac{1-\mu\epsilon^{-2ik\mathrm{L}}}{1-\beta\mu\epsilon^{-2ik\mathrm{L}}}\mathrm{F} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

and

$$\left.\begin{aligned}\mathrm{F}+\beta\mathrm{A}&=\frac{1-\beta}{1-\beta\mu\epsilon^{-2ik\mathrm{L}}}\mathrm{F},\\ \mathrm{F}+\mathrm{A}&=\frac{1-\beta}{1-\beta\mu\epsilon^{-2ik\mathrm{L}}}\mu\epsilon^{-2ik\mathrm{L}}\mathrm{F}.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (6)$$

So that the total potential within the pipe is

$$\begin{aligned}\phi&=\frac{(1-\beta)\mathrm{F}}{1-\beta\mu\epsilon^{-2ik\mathrm{L}}}\{\epsilon^{-ikx}+\mu\epsilon^{-2ik\mathrm{L}}\epsilon^{ikx}\}\epsilon^{ikat}\\ &=\frac{(1-\beta)\mathrm{F}}{1-\beta\mu\epsilon^{-2ik\mathrm{L}}}\{\epsilon^{-ik(x-\mathrm{L})}+\mu\epsilon^{ik(x-\mathrm{L})}\}\epsilon^{ikat}, \quad . \quad . \quad . \quad (7)\end{aligned}$$

* For an account of the recent development of impedance methods in acoustics, see I. B. Crandall's 'Theory of Vibrating Systems and Sound' (Macmillan, 1926).

representing two trains of plane-waves travelling in opposite directions along the axis of the pipe. If the amplitude of the wave travelling from the open to the closed end is unity, the amplitude of the reflected wave is $|\mu|$.

Let $\beta=\beta_0\epsilon^{-\iota k a}$, $\mu=\mu_0\epsilon^{-\iota k\delta}$, where β_0 and μ_0 are real. Then

$$\phi=\frac{(1-\beta)\mathrm{F}}{\epsilon^{\iota k\mathrm{L}}-\beta_0\mu_0\epsilon^{-\iota k(\mathrm{L}+\alpha+\delta)}}\{\epsilon^{-\iota k(x-\mathrm{L})}+\mu\epsilon^{\iota k(x-\mathrm{L})}\}\epsilon^{\iota k a t}$$

$$=\frac{(1-\beta)\mathrm{F}\epsilon^{\frac{1}{2}\iota k(\alpha+\delta)}}{\epsilon^{\iota k\mathrm{L}'}-\beta_0\mu_0\epsilon^{-\iota k\mathrm{L}'}}\{\epsilon^{-\iota k(x-\mathrm{L})}+\mu\epsilon^{\iota k(x-\mathrm{L})}\}\epsilon^{\iota k a t}, \quad . \quad . \quad (8)$$

where $\mathrm{L}'=\mathrm{L}+\frac{1}{2}(\alpha+\delta)$.

To find the condition for resonance, consider the effect of varying L when the wave-length of the sound and the diameter of the pipe are kept constant. The square of the amplitude of the vibration within the pipe will be inversely proportional to the square of the modulus of the denominator of the right-hand member of (8); that is, to

$$|\epsilon^{\iota k\mathrm{L}'}-\beta_0\mu_0\epsilon^{-\iota k\mathrm{L}'}|^2$$

$$=1+\beta_0^2\mu_0^2-2\beta_0\mu_0\cos 2k\mathrm{L}'. \quad . \quad . \quad . \quad (9)$$

This will be a maximum or a minimum when $\sin 2k\mathrm{L}'=0$, or $2k\mathrm{L}'=n\pi$, when n is an integer. Thus when $\mathrm{L}'=n(\lambda/4)$ the energy of the vibration in the pipe * is either a maximum or a minimum, and it is easily seen from (9) that maxima occur when $\mathrm{L}'=0$, $\lambda/2$, λ, $3\lambda/2$, etc., and minima when $\mathrm{L}'=\lambda/4$, $3\lambda/4$, $5\lambda/4$, etc.

The case of a pipe stopped by a complete reflector is arrived at by putting $\Omega=0$, and hence, by (4), $\mu=1$. Thus, for a stopped pipe, the condition for resonance is

$$\mathrm{L}'=\mathrm{L}+\tfrac{1}{2}\alpha=\lambda/2, \quad \lambda, \quad 3\lambda/2, \text{ etc.}$$

The relation of L′ to the customary "reduced length" of a stopped pipe can be seen by considering more closely the meaning of α.

The value of β for a flanged open end is †

$$\beta=-\frac{1-\iota k\sigma\left(\frac{1}{c_0}-\frac{\iota k}{2\pi}\right)}{1+\iota k\sigma\left(\frac{1}{c_0}-\frac{\iota k}{2\pi}\right)}. \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Where c_0 is "the conductance of the open end of the pipe," which, in the case of a circular cross-section, is approximately

* *I. e.*, the average energy per unit length of pipe, not the total energy.

† *Cf.* Phil. Mag. ii. p. 756 (1926).

equal to 3·8 × the radius *. To find α we have

$$\beta = -\frac{\left\{\left(1-\frac{\sigma k^2}{2\pi}\right)-\iota k\sigma/c_0\right\}\left\{1+\frac{\sigma k^2}{2\pi}\right)-\iota k\sigma/c_0\right\}}{\left(1+\frac{\sigma k^2}{2\pi}\right)^2+\left(\frac{\sigma k}{c_0}\right)^2}$$

$$= -\frac{\left\{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2\right\}-2\iota\frac{\sigma k}{c_0}}{\left(1+\left(\frac{\sigma k^2}{2\pi}\right)^2+\left(\frac{\sigma k}{c_0}\right)^2\right.} \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

So that

$$-k\alpha = \tan^{-1}\left\{-\frac{2\sigma k/c_0}{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2}\right\} . \quad . \quad . \quad (12)$$

In the case of a pipe having a diameter small compared with the wave-length, the quantity $\left\{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2\right\}$ is always positive, and hence $\tan(-k\alpha)$ is negative. Thus

$$-k\alpha = \pi - \tan^{-1}\left\{\frac{2\sigma k/c_0}{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2}\right\} . \quad . \quad . \quad (13)$$

and

$$\alpha = \frac{\lambda}{2\pi}\tan^{-1}\left\{\frac{2\sigma k/c_0}{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2}\right\}-\frac{\lambda}{2} . \quad . \quad (14)$$

Let

$$2\alpha' = \frac{\lambda}{2\pi}\tan^{-1}\left\{\frac{2\sigma k/c_0}{\left(1-\frac{\sigma k^2}{2\pi}\right)^2-\left(\frac{\sigma k}{c_0}\right)^2}\right\} . \quad . \quad . \quad (15)$$

Then

$$\alpha = 2\alpha' - \frac{\lambda}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

and

$$L' = L + \tfrac{1}{2}\alpha = L + \alpha' - \frac{\lambda}{4} . \quad . \quad . \quad . \quad . \quad (17)$$

Thus resonance occurs in a narrow stopped pipe when $L+\alpha'-\lambda/4 = 0, \lambda/2, 4\lambda, 3\lambda/2$, etc. ; that is, when $L+\alpha' = \lambda/4$,

* *Ibid.* p. 757.

$3\lambda/4$, $5\lambda/4$, etc., which agrees with the well-known facts if α' is the usual "correction for an open end." The expression for α' in (15) reduces to the approximate relation, $\alpha'=\sigma/c_0$, when the diameter of the pipe is very small compared with the wave-length *.

To return to the case when the pipe is stopped with an admittance Ω at $x=L$, we see that resonance occurs when $L+\alpha'+\frac{1}{2}\delta=\lambda/4, 3\lambda/4, 5\lambda/4$, etc.; that is, when the "reduced length" of the pipe $(L+\alpha')$ plus the distance $\frac{1}{2}\delta$, which depends on the change of phase at reflexion at the stopped end, is equal to an odd number of quarter wave-lengths.

Some special cases may be mentioned. If the pipe is stopped at $x=L$ by a mass of non-porous material, there is no change of phase at reflexion and $\delta=0$, so that the reduced length $(L+\alpha')$ gives greatest resonance no matter how much sound-energy is absorbed at the stopped end. This case is, however, somewhat difficult to realize in practice on account of the necessarily finite thickness of the layer of material used to stop the pipe †.

If the material is porous, there may be change of phase at reflexion, in which case $\delta\neq 0$. The difference in the lengths of a pipe adjusted to resonance when stopped with porous and non-porous materials might, in fact, be used to measure δ and hence the change of phase at reflexion.

As an example of a somewhat different kind, reference may be made to the Boys pattern double resonator‡ (fig. 2). The pipe in this case is closed by a rigid plate in which there is a small aperture forming the entrance to a Helmholtz resonator. The theory of the Helmholtz resonator shows that its admittance is given by

$$\Omega=\frac{c}{2h-\iota\Delta}, \qquad \ldots\ldots \quad (18)$$

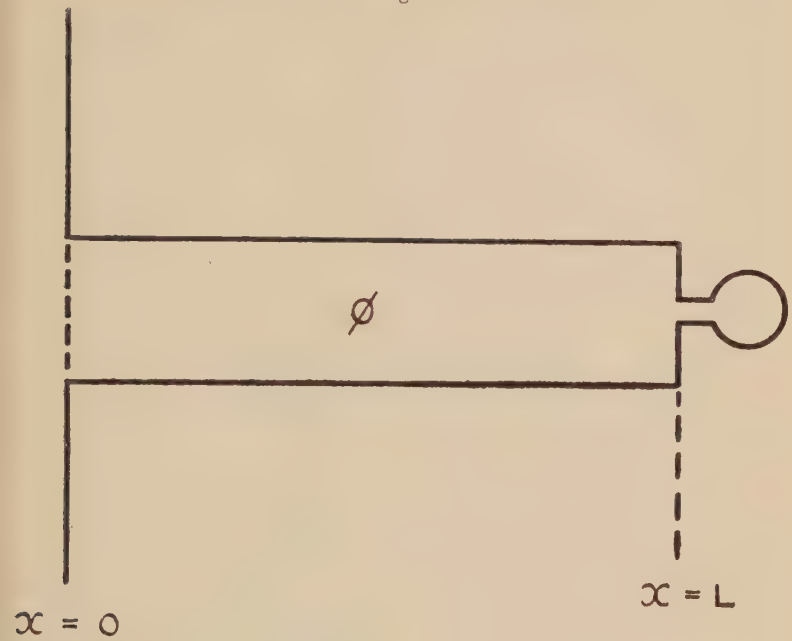
where c is the conductance of the orifice, h is the damping coefficient, and $\Delta=\omega_0(\omega_0/\omega-\omega/\omega_0)$, ω_0 being equal to 2π times the resonance-frequency of the resonator, and ω to 2π times the frequency of the incident sound-waves. Thus in

* *Cf.* Rayleigh, 'Theory of Sound,' ii. (2nd ed.) p. 200.

† If the pipe is stopped by a thin diaphragm the appropriate value of Ω is $(\sigma\rho/m)/(2h-\iota\Delta)$, in which ρ is the density of air, m is the equivalent mass of the diaphragm, h is its damping factor and $\Delta=\omega_0(\omega_0/\omega-\omega/\omega_0)$, where $\omega_0=2\times$ the resonance frequency of the diaphragm and $\omega=2\pi\times$ the frequency of the sound.

‡ The theory of this resonator is dealt with in Phil. Mag. ii. pp. 751–769 (1926).

Fig. 2.



the case of a Boys resonator *

$$\mu=\frac{1-\dfrac{ac}{\sigma(2h-\iota\Delta)}}{1+\dfrac{ac}{\sigma(2h-\iota\Delta)}}. \quad . \quad . \quad . \quad . \quad . \quad (19)$$

Hence

$$\mu_0^2=\frac{\left\{\Delta^2+4h^2-\left(\frac{ac}{\sigma}\right)^2\right\}^2+4\Delta^2\left(\frac{ac}{\sigma}\right)^2}{\left\{\left(2h+\frac{ac}{\sigma}\right)^2+\Delta^2\right\}^2}, \quad . \quad . \quad (20)$$

$$\delta=\frac{\lambda}{4\pi}\tan^{-1}\left\{\frac{2\Delta\left(\frac{ac}{\sigma}\right)}{\Delta^2+4h^2-\left(\frac{ac}{\sigma}\right)^2}\right\}. \quad . \quad . \quad . \quad (21)$$

The expression (20) is interesting, for it shows that if $\Delta=0$, that is if the Helmholtz resonator is exactly in unison with the incident sound, and also if $2h=\dfrac{ac}{\sigma}$, then

* μ has not the same meaning as in the earlier paper. It is used to denote the reciprocal of μ as previously used.

$\mu_0=0$; that is, *the sound-waves passing down the pipe from the open end are completely absorbed at* $x=L$. The admittance at $x=L$ in this case is therefore equal to that of an infinite pipe of cross-section σ. These conditions are not difficult to fulfil. At 512 vibrations per second, for example, it is possible to make a Helmholtz resonator with $h=80$ sec.$^{-1}$ and $c=0{\cdot}125$ cm. Thus for complete absorption we should have $\sigma=\frac{ac}{2h}=\frac{33760\times 0{\cdot}125}{2\times 80}=26{\cdot}4$ cm.^{2} If the pipe is circular $\sigma=\pi R^2$, and therefore $R=\sqrt{26{\cdot}4/\pi}=2{\cdot}9$ cm.

This result is independent of L, the length of the pipe, and appears to be of wider application than to the theory of resonance in pipes. It suggests that by means of a certain distribution of resonators over the face of a flat reflecting wall complete absorption of normally incident sound of a certain wave-length could be secured by a suitable choice of the dimensions of the resonators.

From the expression (21), and the condition that $L+\alpha'+\frac{1}{2}\delta$ must be equal to an odd number of quarter wave-lengths, the values of L which will make the energy of the vibration in the pipe a maximum can be found. If $\Delta=0$, then $\delta=0$, so that if the incident sound and the Helmholtz resonator are in tune, best resonance (for this frequency) is obtained when the reduced length of the pipe is an odd number of quarter wave-lengths.

The case when the source of sound is inside the pipe will now be briefly considered. Let the source lie in the plane $x=0$ and let the strength be $Ue^{\iota kat}$ c.c. per sec., such as might be provided by a piston closely fitting into the tube and driven by suitable mechanism, or by a telephone diaphragm driven by alternating current.

The potential within the pipe is

$$\phi=(Ae^{-\iota kx}+Be^{\iota kx})e^{\iota kat}, \quad . \quad . \quad . \quad . \quad (22)$$

and we must have

$$-\sigma\frac{\partial\phi}{\partial x}=Ue^{\iota kat} \text{ when } x=0, \quad . \quad . \quad . \quad (23)$$

whence

$$B=A-U/\iota k\sigma. \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Also when $x=L$,

$$-\sigma\frac{\partial\phi}{\partial x}=\Omega\frac{\partial\phi}{\partial t}. \quad . \quad . \quad . \quad . \quad . \quad (25)$$

Therefore

$$A - \frac{U}{\iota k\sigma} = \mu A \epsilon^{-2\iota kL}, \quad . \quad . \quad . \quad . \quad . \quad (26)$$

where μ has the same meaning as previously.

Thus

$$\left.\begin{aligned} A &= \frac{U/\iota k\sigma}{1-\mu\epsilon^{-2\iota kL'}} \\ B &= \frac{U/\iota k\sigma}{1-\mu\epsilon^{-2\iota kL}}\mu\epsilon^{-2\iota kL}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (27)$$

So that

$$\begin{aligned} \phi &= \frac{U/\iota k\sigma}{1-\mu\epsilon^{-2\iota kL}}\{\epsilon^{-\iota kx} + \mu\epsilon^{-2\iota kL}\epsilon^{\iota kx}\}\epsilon^{\iota kat} \\ &= \frac{U/\iota k\sigma}{\epsilon^{\iota kL} - \mu\epsilon^{-\iota kL}}\{\epsilon^{-\iota k(x-L)} + \mu\epsilon^{\iota k(x-L)}\}\epsilon^{\iota kat}. \quad . \quad . \quad (28) \end{aligned}$$

If, as before, $\mu = \mu_0\epsilon^{-\iota k\delta}$,

$$\phi = \frac{U/\iota k\sigma}{\epsilon^{\iota kL'} - \mu_0\epsilon^{-\iota kL'}}\{\epsilon^{-\iota k(x-L')} + \mu_0\epsilon^{\iota k(x-L')}\}\epsilon^{\iota kat}, \quad . \quad (29)$$

where $L' = L + \frac{1}{2}\delta$.

The square of the amplitude of the vibration in the pipe is inversely proportional to

$$\begin{aligned} &|\epsilon^{\iota kL'} - \mu_0\epsilon^{-\iota kL'}|^2 \\ &= (1+\mu_0)^2 - 2\mu_0\cos 2kL'. \quad . \quad . \quad . \quad . \quad (30) \end{aligned}$$

The energy of the vibration is therefore a minimum or a maximum when $\sin 2kL' = 0$, or $L' = n\frac{\lambda}{4}$. It is easily seen that resonance occurs, *i. e.*, the energy is a maximum, when n is even, that is when $L' = \frac{\lambda}{2}$, λ, etc.

If the pipe is closed by a rigid plate at $x = L$, $\Omega = 0$ and $\mu = 1$. Hence, by (29), the amplitude of the vibration in the pipe is inversely proportional to $\sin kL'$, and becomes (theoretically) infinite when $\sin kL' = 0$; that is, when $L' = \frac{\lambda}{2}$, λ, etc. This may be taken as indicating that the amplitudes become so great that the approximate equations of sound-propagation on which the theory given above is founded are no longer applicable.

Equations (7) and (28) find an application in connexion with the theory of the stationary-wave method of measuring sound-absorption at normal incidence. The principle of this

method can be illustrated by considering the case of plane-waves incident normally on the flat surface of sound-absorbing material. Let the reflecting surface lie in the plane $x=0$, and μ be the coefficient of reflexion. Then the amplitude of the incident waves being supposed equal to unity, the potential of the incident and reflected waves will be

$$\phi=(\epsilon^{-\iota kx}+\mu\epsilon^{\iota kx})\epsilon^{\iota kat}. \quad . \quad . \quad . \quad . \quad (31)$$

It is clear that this potential will give rise to a series of positions of maximum and minimum pressure-variation in front of the reflecting surface, and that (if $\mu=\mu_0\epsilon^{-\iota k\delta}$) the ratio of maximum and minimum will be $(1+\mu_0)/(1-\mu_0)$. If this ratio is measured with an acoustical instrument and is found to be, say, m, then $\mu_0=(m-1)/(m+1)$. The proportion of the incident sound-energy which if absorbed, namely $(1-\mu_0^2)$, can therefore be determined in terms of the observed ratio m. If Ω' is the acoustical admittance per unit area of the reflecting surface at normal incidence, then, since ϕ must satisfy the condition

$$-\frac{\partial\phi}{\partial x}=\Omega'\frac{\partial\phi}{\partial t} \text{ when } x=0, \quad . \quad . \quad . \quad . \quad (32)$$

we find that $\mu=(1-a\Omega')/(1+a\Omega')$.

In practice, in order to produce plane-waves incident normally on a reflecting surface, recourse is had to a pipe stopped at one end with a disk of the material to be tested. The other end of the pipe may be left open and a source may be placed near it, or, alternatively, the source may be actually mounted in the other end of the pipe. The two arrangements correspond to the two cases considered above, and the resulting potential inside the pipe is given by equations (7) and (28) respectively. It will be seen from these equations that the form of the potential inside the pipe is the same in both cases, and depends only on the value of the admittance placed at $x=\mathrm{L}$. The potential is that due to two sets of plane-waves travelling in opposite directions; and if the amplitude of that travelling in the direction x-positive is unity, then the amplitude of that travelling in the direction x-negative is μ_0, $\mu(=\mu_0\epsilon^{-\iota k\delta})$ being the coefficient of reflexion at $x=\mathrm{L}$. There are thus positions of maximum and minimum pressure-amplitude inside the pipe, just as there are in front of an infinite plane-reflecting face. Also we see from (4) that $\mu=(1-a\Omega')/(1+a\Omega')$, where Ω' is written for Ω/σ and is the admittance per unit area of the reflecting surface. Thus μ has the same meaning as in (31).

It may be noted that the ratio of maximum to minimum pressure-amplitude inside the pipe (though not, of course, the amplitudes themselves) is independent of L, the length of the pipe. This point appears to have been the subject of a certain amount of misconception. Eckhardt and Chrisler *, for example, in recent work with the stationary-wave method, brought the pipe into resonance with the source at each frequency at which observations were made, appearing to think that the measurements would otherwise be unreliable. Also Tuma †, who was the first to use the method, brought the pipe into resonance before making observations. Such a procedure seems to be unnecessary.

Summary.

The theory of pipes stopped with imperfect reflectors is discussed and the conditions for resonance are deduced. Some special cases are dealt with, including that of a Boys pattern double resonator, and the conditions necessary for complete absorption (*i. e.* no reflexion) of sound by a resonator placed at the end of a stopped pipe are found. Some of the results obtained have an application to the theory of the measurement of sound-absorption by the stationary-wave method.

Biggin Hill, Kent.
June 1927.

LXXXVI. *The Effect of Variable Specific Heats upon the Velocity Generated, and upon the Temperature Drop, in Gases Expanding through Nozzles.* *By* E. C. Wadlow, *Ph.D., B.Sc.* ‡

EQUATIONS for calculating the velocity generated, and the weight of gas discharged, when a gas initially at a high pressure is allowed to expand through a nozzle to a lower pressure, taking into consideration the increase of specific heat with temperature, have been given by Dr. W. J. Walker §.

* 'Scientific Papers of the Bureau of Standards,' No. 526 (April, 1926).

† *Wien. Sitzungber. d. K. Akad. d. Wissenschaften,* iii. pt. 2A, pp. 402–410 (1902).

‡ Communicated by the Author.

§ Phil. Mag. xliii. pp. 589–593 (1922).

The formulæ giving the velocity generated are based upon the fundamental equation

$$V^2 = 2gK_p \int_{T_2}^{T_1} dT, \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where

V = Velocity generated. Feet per second,
g = 32·2 feet per sec. per sec.,
K_p = Specific heat of gas at constant pressure. Ft. lb. per lb. per ° C.
T_1 = Initial temperature ° C. abs.
T_2 = Final temperature ° C. abs.

If the relationship between specific heat and temperature is given by

$$\left.\begin{aligned} K_p &= A + ST \\ K_v &= B + ST \end{aligned}\right\}, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

the relationship between velocity generated and temperature is given by

$$V^2 = 2g\left[A(T_1 - T_2) + \frac{S}{2}(T_1^2 - T_2^2)\right]. \quad . \quad . \quad (3)$$

Similarly, if

$$\left.\begin{aligned} K_p &= A + ST + S'T^2 \\ K_v &= B + ST + S'T^2 \end{aligned}\right\}, \quad . \quad . \quad . \quad . \quad . \quad (4)$$

the velocity generated is given by

$$V^2 = 2g\left[A(T_1 - T_2) + \frac{S}{2}(T_1^2 - T_2^2) + \frac{S'}{3}(T_1^3 - T_2^3)\right]. \quad . \quad (5)$$

Owing to the difficulty of measuring T_2 when the velocity of the gas issuing from the nozzle is high, equations (3) and (5) cannot be used directly to calculate, from experimental observations, the actual velocity generated during the expansion, although the numerical values of g, A, S, S′ and T_1 may be known or measured with fair accuracy.

With a view to overcoming this difficulty, equations (3) and (5) have been transformed by Dr. Walker to the following forms, respectively:

$$V = \left[2g\frac{m}{m-1}p_1v_1\left\{1-\left(\frac{p_0}{p_1}\right)^{\frac{m-1}{m}}\right\}\right]^{\frac{1}{2}} \left\{1+\frac{\lambda T_1}{4m}\left(1-\left(\frac{p_0}{p_1}\right)^{\frac{m-1}{m}}\right)\right\}. \quad . \quad (6)$$

and

$$V=\left[2g\frac{m}{m-1}p_1v_1\left\{1-\left(\frac{p_0}{p_1}\right)^{\frac{m-1}{m}}\right\}\right]^{\frac{1}{2}}$$

$$\left\{1+\left(\lambda T_1+\frac{\lambda'}{2}T_1^2\right)\left[\frac{1-\left(\frac{p_0}{p_1}\right)^{\frac{m-1}{m}}}{4m}\right]\right\}. \quad . \quad (7)$$

where

$$m=\frac{A}{B}, \qquad \lambda=\frac{S}{B}, \qquad \lambda'=\frac{S'}{B}.$$

p_1 = Initial pressure. Lb. per square foot.
v_1 = Initial specific volume. Cubic ft. per lb.
p_0 = Final pressure. Lb. per square foot.
T_1 = Initial temperature, ° C. abs.

Thus the velocity may be calculated from observations of the initial and final pressures, initial temperature, and physical constants for the gas under consideration, all of which may be measured fairly accurately without difficulty.

For many purposes, however, a knowledge of the theoretical temperature of the gas at the end of the expansion is desirable, and this is particularly so when this quantity is to be used as a standard for comparison with values obtained by measurement or by indirect methods. A knowledge of this temperature also enables equations (3) and (5) to be employed, without further transformation, for calculating the theoretical velocity generated.

It is the object of the following to supplement the equations of Dr. Walker by giving formulæ for the calculation of the temperature of the gas at the end of an adiabatic expansion, and thus to provide an alternative method of calculating the theoretical velocity generated.

The most general case is obtained by assuming the relationship between specific heats and temperature of the gas to be given by (4). The adiabatic equation may then be written

$$pv^m e^{\lambda T+\frac{\lambda'}{2}T^2}=\text{constant}, \quad . \quad . \quad . \quad . \quad (8)$$

where the symbols have the meanings previously assigned to them, and in addition, e is the base of natural logarithms. Therefore

$$p_1{v_1}^m e^{\lambda T_1+\frac{\lambda'}{2}T_1^2}=p_2{v_2}^m e^{\lambda T_2+\frac{\lambda'}{2}T_2^2}. \quad . \quad . \quad (9)$$

Also $$pv = \mathrm{RT}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Putting $$v = \frac{\mathrm{RT}}{p}$$

in (9) and simplifying,

$$p_1^{1-m}\mathrm{T}_1^{m}e^{\lambda \mathrm{T}_1+\frac{\lambda'}{2}\mathrm{T}_1^2} = p_2^{1-m}\mathrm{T}_2^{m}e^{\lambda \mathrm{T}_2+\frac{\lambda'}{2}\mathrm{T}_2^2}, \quad . \quad . \quad (11)$$

or

$$\left(\frac{p_2}{p_1}\right)^{\frac{1-m}{m}} = \frac{\mathrm{T}_1}{\mathrm{T}_2}e^{\frac{\lambda}{m}(\mathrm{T}_1-\mathrm{T}_2)+\frac{\lambda'}{2m}(\mathrm{T}_1^2-\mathrm{T}_2^2)} \quad . \quad . \quad . \quad (12)$$

Expanding the exponential, and writing x for $\left(\frac{p_2}{p_1}\right)^{\frac{1-m}{m}}$ or $\left(\frac{p_1}{p_2}\right)^{\frac{m-1}{m}}$, we have

$$\mathrm{T}_2 x = \mathrm{T}_1\Big[1+\frac{\lambda}{m}(\mathrm{T}_1-\mathrm{T}_2)+\frac{\lambda'}{2m}(\mathrm{T}_1^2-\mathrm{T}_2^2) + \Big\{\frac{\lambda}{m}(\mathrm{T}_1-\mathrm{T}_2)\Big\}^2+\ldots\Big]. \quad . \quad (13)$$

Neglecting terms involving squares and higher powers of $\left(\frac{\lambda}{m}\right)$ and of $\left(\frac{\lambda'}{m}\right)$, we have

$$\mathrm{T}_2 x = \mathrm{T}_1+\frac{\lambda}{m}\mathrm{T}_1^2-\frac{\lambda}{m}\mathrm{T}_1\mathrm{T}_2+\frac{\lambda'}{2m}\mathrm{T}_1^3-\frac{\lambda'}{2m}\mathrm{T}_1\mathrm{T}_2^2, \quad . \quad (14)$$

or

$$\mathrm{T}_2\Big(x+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2m}\mathrm{T}_1\mathrm{T}_2\Big)=\mathrm{T}_1\Big(1+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2m}\mathrm{T}_1^2\Big), \quad . \quad (15)$$

or

$$\mathrm{T}_2=\mathrm{T}_1\left[\frac{1+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2m}\mathrm{T}_1^2}{x+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2m}\mathrm{T}_1\mathrm{T}_2}\right], \quad . \quad . \quad . \quad (16)$$

or, since $\mathrm{T}_2=\frac{\mathrm{T}_1}{x}$ nearly,

$$\mathrm{T}_2=\mathrm{T}_1\left[\frac{1+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2m}\mathrm{T}_1^2}{x+\frac{\lambda}{m}\mathrm{T}_1+\frac{\lambda'}{2mx}\mathrm{T}_1^2}\right]. \quad . \quad . \quad . \quad . \quad (17)$$

A more convenient and, for most purposes, sufficiently

accurate formula is obtained by neglecting all terms involving λ', and terms involving squares and higher powers of $\left(\frac{2}{m}\right)$.

We then have

$$T_2 = T_1\left[\frac{1+\frac{\lambda}{m}T_1}{x+\frac{\lambda}{m}T_1}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

Equations (17) and (18) may be simplified by writing

$$\alpha = \frac{\lambda}{m}, \qquad \beta = \frac{\lambda'}{2m},$$

when they become

$$T_2 = T_1\left[\frac{1+\alpha T_1+\beta T_1^2}{x+\alpha T_1+\frac{\beta}{x}T_1^2}\right] \quad . \quad . \quad . \quad . \quad (17\,a)$$

and

$$T_2 = T_1\left[\frac{1+\alpha T_1}{x+\alpha T_1}\right] \quad . \quad . \quad . \quad . \quad . \quad (18\,a)$$

respectively.

The values of α and β for a number of gases which find application industrially have been calculated from data given by Partington and Shilling *. These hold for a temperature range of 0° C. to 2000 ° C., and are given in Table I.

TABLE I.

Gas.	α.	β.
Air	0·000,024,5	0·000,000,022,4
Oxygen	0·000,024,5	0·000,000,022,4
Nitrogen	0·000,024,5	0·000,000,022,4
Carbon Monoxide ...	0·000,024,5	0·000,000,022,4
Carbon Dioxide	0·000,592	−0·000,000,067
Hydrogen	0·000,105	zero.
Steam	−0·000,132	0·000,000,129

The velocity generated may now be obtained either by substituting T_2 from (17 *a*) in (5) or T_2 from (18 *a*) in (3). For most purposes it is sufficiently accurate to use (18 *a*) and (3), and decidedly more convenient arithmetically.

* 'The Specific Heats of Gases,' pp. 201–209. Benn Bros., Ltd.

Regarding the calculation of the velocity generated during adiabatic expansion, with variable specific heats, there is little to chose between the equations of Dr. Walker and the method employed here. Both require the same number of initial measurements and the accuracy of both is dependent upon the accuracy with which the numerical values of the relevant properties of the gases are known.

The equations are likely to be most useful in connexion with the design of internal combustion turbines. The maximum pressures and temperatures which can be foreseen are in the neighbourhood of 300 lb. per square inch and 2500° C. abs., respectively. The field is such, therefore, that changes of specific heat with temperature are all important, and changes with pressure will be so small as to be safely neglected.

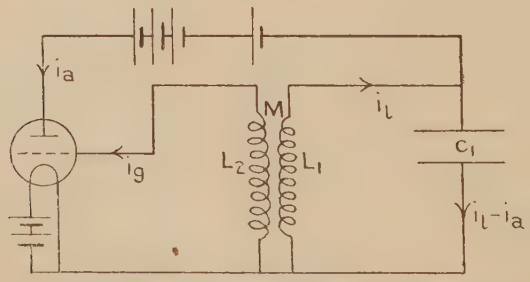
LXXXVII. *On Frequency Variations of the Triode Oscillator.*

By David F. Martyn, *B.Sc., A.R.C.Sc., Research Student, University of Glasgow* *.

(1) *Introduction.*

IT is well known that the frequency of the oscillations generated by the simple oscillator of fig. 1 does not depend solely on the values of the inductance L_1 and the capacity C_1. Thus small changes, of the order of 2 or 3 per cent., may be produced by alteration of the filament

Fig. 1.



current or anode voltage of the valve, or by altering the coupling between the coils L_1 and L_2.

These variations have been examined experimentally by several investigators, notably by Eccles and Vincent †. So

* Communicated by Professor E. Taylor Jones, D.Sc.

† Proc. Roy. Soc. xcvi. and xcvii. (1920).

far no satisfactory explanation of the phenomena has been offered *.

The purpose of the present paper is to record various new effects which have been obtained experimentally, and to put forward a theory which appears to be capable of explaining the observed phenomena.

(2) *Apparatus.*

In all the experiments to be described below the simple oscillator of fig. 1 was used. Two different pairs of generating coils were used. One pair had each a coefficient of self induction of 0·03125 henries. The coefficient of mutual induction between the coils of this pair could be varied from 0·02182 henries downwards. This pair of coils will hereafter be denoted as Coils I. The other pair of coils were considerably larger. One of this pair (A) had a coefficient of self induction of 0·665 henries, while that of the other (B) was 1·49 henries. The mutual inductance between these coils was fixed and had a value of 0·752 henries. This pair will be denoted as Coils II.

The frequency variations normally obtained with Coils I. were of the order of magnitude noticed by other observers ; *i.e.*, about 3 per cent. They were detected by a heterodyne method, a second oscillator being loosely coupled to the main oscillator, and the changes in the beat frequency observed.

In the case of Coils II. very large frequency changes were obtained in the circumstances to be described later. These changes extended sometimes over a range of three octaves. In these cases, the frequency being audible, the changes were measured by comparison with a monochord and standard tuning-fork.

(3) *Variation of Frequency with Filament Current.*

Using Coils II. and a Mullard P.M. 6 valve, the following results were obtained (fig. 2):—In the figure the ordinate represents frequency in cycles per second, and the abscissa filament current in milliamperes. The oscillating circuit L_1C_1 consisted of the coil (A) in parallel with a condenser of capacity 0·03 microfarad.

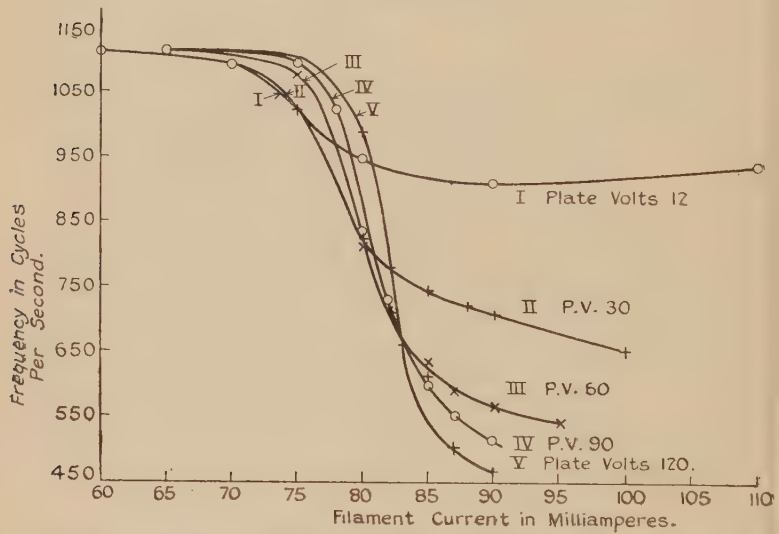
It is seen from the graph that in each case when oscillations commence they have very nearly the natural frequency $\frac{1}{2\pi\sqrt{L_1C_1}}$. The general effect of increase of filament

* See, however, note at end of paper.

current is to produce a decrease of frequency. This decrease is greater for high values of plate voltage. For very low plate voltages a minimum frequency is obtained. This minimum occurs at a high value of the filament current.

The initial rate of decrease of frequency is slower for high than for low plate voltages. As the filament current increases, however, a point is soon reached at which the rate of decrease of frequency for high plate voltage becomes much greater than that for low plate voltages. As the filament current reaches its maximum, the curves tend to straighten out, the rate of change of frequency with filament current again becoming small.

Fig. 2.



The variations of frequency extend over more than two octaves for high plate voltages.

Similar types of curves were obtained, using valves of very different characteristics.

Using Coils I., similar types of variation were observed, but the total frequency changes were smaller in magnitude, being, say, 3 per cent. when the same capacity (0·03 mfd.) as formerly was used. Since these coils were movable, it was possible to examine the effect of variation of coupling on the character of the frequency-filament current curves.

It was found that if loose coupling was used, then the minimum frequency could be obtained at higher values of

the plate voltage than was possible when using close coupling. Further, the minimum was now obtained for a lower value of filament current, and the total range of the frequency variations was reduced.

The attainment of this minimum frequency is important from the practical point of view. Eccles and Vincent * have pointed out that, in order to obtain a constant frequency from a triode oscillator, it is desirable to set the filament current to the value which gives the minimum frequency. Small changes in filament current will then produce the smallest possible change of frequency. Eccles and Vincent indicate that this minimum frequency can always be obtained by the use of a weak enough coupling. On the other hand, recently Obata † has invariably failed to obtain the minimum frequency, although using several different types of valve. The results of the writer, as described above, indicate that the value of the plate voltage plays an important part in determining the position of the minimum frequency. Thus it is often possible to obtain the minimum while using the closest coupling, provided that the plate voltage is sufficiently low. On the other hand, when using loose coupling, the minimum is not obtained if too high a value of plate voltage is used. The conditions necessary for the production of a minimum frequency are further examined in the sequel, when the theoretical explanation of the frequency changes is considered.

The variation of the oscillatory current which occurred in the coil L_1 when the filament current was altered was also examined. It was found that normally the effect of increase of filament current was to produce an increase of the oscillatory current. Whenever a minimum occurred on the frequency-filament current curve, however, there simultaneously occurred a minimum on the curve connecting oscillatory current and filament current. This effect has been previously noted by Vincent ‡. It was further observed that a minimum occurred at the same place on the curve connecting oscillatory grid current and filament current.

After a valve has been run for some time at a high filament current, then in some cases the minimum frequency on the frequency-filament current curve is no longer obtainable. In general, the older a valve is, and the longer it has been in use, the more difficult it is to obtain the minimum frequency.

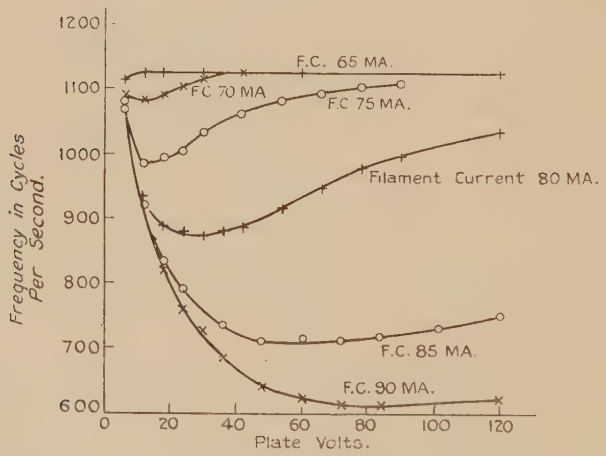
* *Loc. cit.*
† Proc. Phys.-Math. Soc. Japan, Jan. 1927.
‡ *Loc. cit.*

(4) *Variation of Frequency with Plate Voltage.*

Typical curves illustrating the variation of frequency with plate voltage are given in fig. 3. These results were obtained with Coils II., the condenser of capacity 0·03 mfd. being connected across coil A. The valve used was a Mullard P.M. 6. Curves are plotted for values of the filament current ranging from 65 to 90 milliamperes.

As the plate voltage is increased from zero, oscillations commence at about six volts with a frequency slightly below the value $\frac{1}{2\pi\sqrt{L_1C_1}}$. Further increase of plate voltage

Fig. 3.



produces a decrease in frequency. At a certain value of plate voltage a minimum frequency occurs, and further increase of voltage produces a rise in frequency. The total frequency change, which may amount to as much as two octaves, is greater for higher values of the filament current. A further effect of using a higher filament current is to move the position of minimum frequency to the right of the graph, so that the minimum occurs at a higher value of plate voltage. For very low filament current values the curve shows an initial rise. In these cases it appears that the minimum has passed off the scale to the left. It is no longer obtainable because oscillations cannot be made to occur at a sufficiently low plate voltage.

Similar types of curves were obtained using quite different makes of valves.

The experiments were repeated using Coils I. The results were of the same character as those obtained with Coils II. ; but the frequency variations were of a smaller order of magnitude, being at most 2 per cent. when using the same value of capacity C_1.

The effect of variation of coupling on the character of the frequency-plate voltage curves was investigated. It was found in every case that the effect of loosening the coupling was to reduce the rate at which the frequency changed with respect to plate voltage. Further, with looser coupling, in every case the position of minimum frequency on the frequency-plate volts curve was shifted to the right : *i. e.*, with looser coupling the minimum frequency occurred at a higher value of the plate potential.

The amplitude of the oscillatory currents in L_1 and L_2 was found to increase steadily as the plate voltage was increased. No minimum amplitude occurred at any point on the curve.

(5) *Conditions Favourable to Greatest Frequency Variation.*

In all the experiments which were conducted on frequency variations, whether produced by change of filament current or of plate voltage, it was observed that the largest percentage changes of frequency were obtained in the following conditions :—

(1) When the capacity of the condenser C_1 was as small as possible. For example, with Coils I. the total possible frequency change could be increased from 3 per cent. to about 50 per cent. by reduction of the capacity C_1 from 0·03 mfd. to 0·002 mfd.
(2) When the coupling between the coils L_1 and L_2 was as close as possible.
(3) When the self inductance of the grid coil L_2 was as large as possible.
(4) When the "impedance" of the valve used was as low as possible.

(6) *Variation of Frequency with Coupling.*

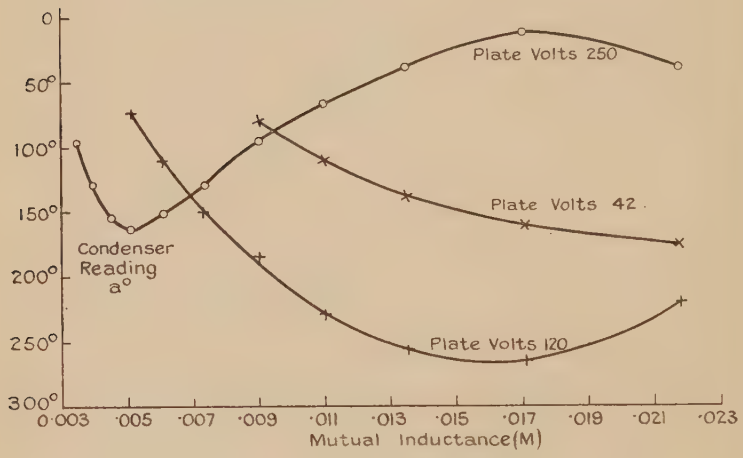
For the purpose of examining the effect of variation of coupling on frequency, Coils I. were used. The condenser C_1 had a capacity of 0·04 mfd. The frequency changes were measured by the beat method. A small variable condenser included in the circuit of the standard oscillator was adjusted so as to bring the standard oscillator into unison with the main oscillator. An increase in the dial reading $a°$

of this condenser corresponds to a decrease in frequency of the oscillator under consideration.

Fig. 4 shows the types of frequency variation obtained. The ordinate represents the condenser reading $a°$, and the abscissa the value of the mutual inductance between the coils. Curves are given for different values of the plate voltage.

Three distinct types of frequency variation occur, according to the value of the plate voltage.

Fig. 4.



For low values of plate voltage the frequency steadily decreases as M increases (type *a*.)

For medium values of plate voltage a minimum frequency occurs (type *b*).

For high values of plate voltage a minimum frequency occurs at a low value of M, and a maximum frequency also occurs at a high value of M (type *c*).

In one instance only, a second minimum was observed at a very high value of M (type c^1). In all other respects this case conformed to type *c*.

The total range of frequency variation is about 2 per cent. The above types of curves were obtained with several different valves. In each case the filament current used was the normal value for the valve in question. Similar curves were obtained at higher and lower values of filament current, provided that the values of the plate voltage were altered to correspond. Thus, when using lower values of

filament current, the typical curves *a*, *b*, and *c* were obtained for lower plate voltages than were necessary in the case considered.

(7) *Theory of Frequency Variations.—Introductory.*

In the generally accepted theory of the production of triode oscillations by the arrangement shown in fig. 1, the following assumptions are made :—

(1) That no self, mutual, or inter-electrode capacities are present, the only capacity considered being that of the condenser C_1.

(2) That no current of electrons flows in the valve between the filament and the grid at any time.

(3) That the anode current i_a is given by the simple expression—

$$i_a = k_1 v_a + k_2 v_g,$$

where v_a and v_g are the anode and grid potentials respectively at any instant, and k_1 and k_2 are constants for a definite value of the filament current.

If, further, for the moment, we neglect the resistances present in the circuit, then it is easy to show that undamped oscillations of frequency $\frac{1}{2\pi\sqrt{L_1C_1}}$ will be produced, provided :

$$k_2 M = k_1 L_1,$$

the latter relation being known as the "condition of maintenance." On this theory, the frequency of the oscillations should depend solely on the values of L_1 and C_1, and no variation should be produced by alteration of the conditions of working of the valve. The theory has been extended by Eccles* to the case where resistances are included in series with both L_1 and C_1.

If R denotes the resistance in series with L_1, and S that in series with C_1, then he finds that the pulsatance "p" of the oscillations is given by the expression :

$$p^2 = \frac{1 + k_1 S}{L_1 C_1 - R C_1^2 (R + S + k_1 R S)},$$

while $$k_2 M = k_1 L_1 + C_1 (k_1 R S + R + S)$$

is the condition of maintenance.

The above expression for "p" contains the quantity k_1. The value of k_1 is dependent on the values of the filament

* Proc. Phys. Soc. xxxi. p. 137 (1919).

current and plate voltage of the valve. Hence we should expect changes in frequency to be produced by alterations in the values of these latter quantities. k_1 has usually a value of about 10^{-5} mho, however, so that the magnitude of the changes will be very small unless the resistances are large. In the case of the apparatus used by the writer the resistances were very small. Thus S was sensibly zero, and R not more than 5 ohms for Coils I. The observed frequency variation was, say, 3 per cent., or about 500 times greater than that possible by the above formula.

A further advance in the theory of the oscillations has been made by Appleton and Greaves *. They expressed the anode current i_a in the form of a power series in v_a by consideration of the "oscillation characteristic" of the valve. As a result of this procedure they succeeded in taking account of the curvature of the valve characteristics, and of the harmonics thereby produced. The presence of these harmonics affected the frequency of the fundamental, and an expression for the fundamental frequency in the form of a series was obtained. Here again, however, the frequency variation accounted for by the theory is of a much smaller order of magnitude than that observed practically.

It appears, therefore, that we must take still other factors into account if we wish to explain the observed frequency variations. All of the experiments described above were carried out at audio-frequencies. Hence we should expect the very small inter-electrode and stray capacities to have no appreciable influence on the frequency in these cases.

There still remains to be considered the question of grid current. It has been customary in theoretical analyses of the problem to assume that this current is infinitesimally small, probably because of the great complication of the analysis which otherwise ensues. When the question of grid current is considered, however, even in its simplest aspect, it appears at once that we have here a factor capable of accounting for a frequency variation of the required order of magnitude. Before considering the theory of the production of oscillations, taking account of grid current, some experimental observations on grid current are given.

(8) *Observations on Grid Current.*

First, Coils II. were used, with a condenser of 0·04 mfd. across Coil A. A current-measuring instrument was inserted between the grid coil and the grid terminal of the

* Phil. Mag. xlv. p. 401 (1923).

valve. The filament current was then increased until oscillations commenced. At once a grid current of twenty microamperes was observed. As the filament current increased, this current increased. As soon as the frequency began to fall appreciably, the grid current commenced to increase more rapidly, until when the frequency had fallen two octaves, the grid current was as much as ten milliamperes.

It has recently been stated * that it is always possible (*e. g.*, by the use of negative grid bias) to make grid current zero while oscillations are occurring. If this statement is correct, it should be possible by this means to attain a very constant frequency, by reason of the elimination of what we are considering to be the chief source of frequency variation; viz., grid current. Accordingly it is of importance to test this statement strictly.

Using Coils II., it was not possible in any circumstances to arrange that the grid-current reading was zero at any time when oscillations were occurring.

With Coils I. the coupling was reduced to the lowest value which permitted of the oscillations being maintained. At this point the oscillations could be heard only faintly, and a delicate galvanometer in the grid circuit registered zero. The question at once arose whether actually no grid current flowed, or whether the galvanometer was not sensitive enough to detect the minute current which was flowing. To test this point, a well-insulated condenser of capacity 2 mfd. was connected in series between the galvanometer and the grid of the valve. The coupling was adjusted, as before, so that oscillations were just maintained, and the galvanometer reading was zero. The oscillations continued for a period of about a minute, and then suddenly died out. If, now, the terminals of the condenser were electrically connected momentarily, then oscillations recommenced, only to die out once more after a lapse of a minute. When a condenser of smaller capacity than 2 mfd. was used, the oscillations died out in a correspondingly shorter interval of time.

It was evident that a very minute grid current was always flowing. This current charged the condenser gradually until the grid received a negative bias large enough to stop the oscillations. It is further evident that if at any instant a negative bias was attained great enough to stop grid current, and yet permitting oscillations to occur, then the

* E. B. Moullin, 'Radio-Frequency Measurements,' p. 7.

bias on the grid would remain constant and the oscillations would continue indefinitely.

We may therefore conclude that it does not appear to be possible for oscillations to occur without flow of current to the grid.

(9) *Theory of Generation of Oscillations, taking account of Grid Current.*

In fig. 1, i_g represents the value of the grid current flowing at any instant, while i_l is the instantaneous value of the current in L_1. If we neglect for the moment the resistances present, then we may write down at once the following equations of the circuit :—

$$v_a = M\frac{di_g}{dt} - L_1\frac{di_l}{dt}, \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$v_g = M\frac{di_l}{dt} - L_2\frac{di_g}{dt}, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$i_l - i_a = C^1\frac{dv_a}{dt}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

As a first approximation, we shall assume that i_a and i_g are each linear functions of v_a and v_g. Hence we have :—

$$i_a = k_1v_a + k_2v_g, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$i_g = k_3v_g + k_4v_a. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

By elimination between these five equations, we obtain the following differential equation giving i_l :—

$$k_3C_1(L_1L_2 - M^2)\frac{d^3i_l}{dt^3} + \{(L_1L_2 - M^2)(k_1k_3 - k_2k_4) + L_1C_1\}\frac{d^2i_l}{dt^2}$$

$$+ \left(L_1k_1 + L_2k_3 - M\,.\,\overline{k_2 + k_4}\right)\frac{di_l}{dt} + i_l = 0. \quad . \quad (6)$$

k_4 is usually very small compared with k_3, and will be neglected to simplify the algebra.

If undamped sinoidal oscillations are produced, then we may write :

$$i_l = I_l \sin pt.$$

Substituting this value of i_l in equation (6) and equating to zero the coefficients of $\sin pt$ and $\cos pt$, we obtain two equations for p^2 :

$$p^2 = \frac{L_1k_1 + L_2k_3 - Mk_2}{k_3C_1(L_1L_2 - M^2)} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

$$= \frac{1}{L_1C_1 + k_1k_3(L_1L_2 - M^2)}. \quad . \quad . \quad . \quad (8)$$

The condition of maintenance is necessarily contained in these two equations.

In addition, another solution of equation (6) exists, of the form :

$$i_l = Ae^{-\alpha t} :$$

α is a very large quantity, having a value normally of, say, 10^5 sec.$^{-1}$. Hence this solution gives rise to a transient which dies out almost instantaneously when oscillations commence, and will therefore be neglected.

It is evident from equation (8) that the frequency of the oscillations generated is less than the natural frequency of the circuit $\frac{1}{2\pi\sqrt{L_1C_1}}$. The amount of this lowering is dependent upon the magnitude of the expression

$$k_1k_3(L_1L_2 - M^2).$$

If we take the approximate numerical value of this expression for an extreme case, using Coils II. with a large grid current flowing, we find that it has a value equal to about four times that of L_1C_1. This indicates a fall of frequency of about two octaves, which is precisely the order of the maximum observed frequency change.

If, now, keeping the same values for k_1, k_3, and C_1, we substitute the values for L_1, L_2, and M appropriate for Coils I., we find that the maximum variation of frequency should be about 3 per cent. This, again, is the order of the maximum frequency variation observed with Coils I.

The quantity k_3, which is difficult to measure experimentally, may be eliminated from equation (8) by the aid of equation (7). Thus

$$k_1k_3(L_1L_2 - M^2)$$

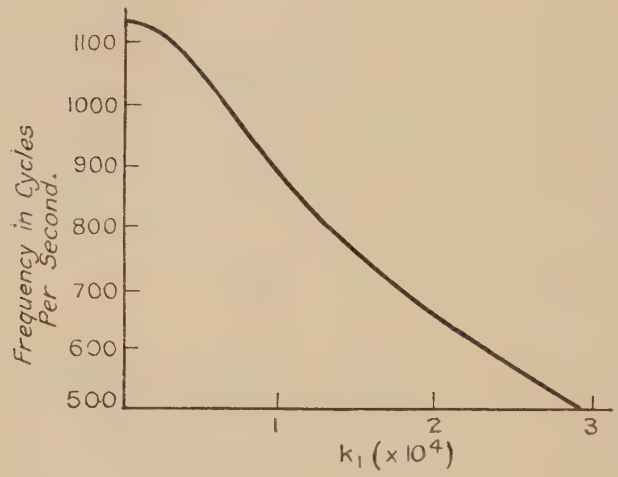
$$= \frac{1}{2L_2}\left\{ -C_1M^2 + k_1^2(L_1L_2 - M^2)(\mu M - L_1) + \sqrt{\begin{matrix} C_1^2M^4 + k_1^4(L_1L_2 - M^2)^2(\mu M - L_1)^2 \\ + 2C_1k_1^2(L_1L_2 - M^2)(\mu M - L_1)(2L_1L_2 - M^2) \end{matrix}} \right\}. \quad . \quad . \quad . \quad (9)$$

In this expression k_2 has been replaced by μk_1, where μ is the "amplification factor" of the valve, a quantity which is appreciably constant over a wide range.

This gives now an expression for the frequency in terms of L_1, L_2, M, C_1, the constant of the valve μ, and k_1. Hence

in any particular case we may plot a curve showing the relation between frequency and k_1. Fig. 5 illustrates the curve so obtained for Coils II. and a Mullard P.M. 6 valve.

Fig. 5.



(10) *Condition of Maintenance.*

Elimination of "p" from equations (7) and (8) leads to the equation :

$$-k_3C_1M^2 = L_1^2k_1C_1 - ML_1C_1k_2 + k_1k_3(L_1L_2 - M^2)(k_1L_1 + k_3L_2 - k_2M). \quad . \quad . \quad . \quad (10)$$

This is the condition for the maintenance of the oscillations. For cases where the variations of frequency are small this reduces to

$$k_3M^2 = k_1L_1(\mu M - L_1). \quad . \quad . \quad . \quad . \quad (11)$$

It is of interest to compare this condition with that obtained on the assumption that grid current is zero—viz. :

$$k_1(\mu M - L_1) = 0.$$

It is difficult to see how this latter equation could ever be satisfied in ordinary circumstances. Thus, in any common case, M might have a value one-quarter that of L_1, and μ might be, say, ten. μM is then considerably greater than L_1. We have to imagine that the oscillations keep on

increasing until μ has decreased greatly to a value which will satisfy the relationship $\mu M = L_1$. Experience indicates that μ is sensibly constant, even over the bends of the valve characteristics, where k_1 and k_2 have changed greatly in value. On the other hand, if we consider equation (11) it is easy to see the course of events. When oscillations commence, the grid current increases until k_3 has a value such that $\frac{k_3 M^2}{k_1 L_1}$ is numerically equal to the difference between μM and L_1.

(11) *Conditions for Maximum Frequency Variation.*

Maximum frequency changes will be obtained when the ratio $\frac{k_1 k_3 (L_1 L_2 - M^2)}{L_1 C_1}$ is as large as possible. From equation (9) we see at once that this condition requires that C_1 shall be as small as possible. When C_1 is small, we have:

$$\frac{k_1 k_3 (L_1 L_2 - M^2)}{L_1 C_1} = \frac{k_1^2}{C_1}\left(1 - \frac{M^2}{L_1 L_2}\right)(\mu M - L_1). \quad . \quad (12)$$

The quantity $\frac{M^2}{L_1 L_2}$ is always less than unity, and it may be kept small by making L_2 large. Hence we see from equation (12) that the largest frequency variations will be obtained when

(1) C_1 is as small as possible,
(2) M and L_2 are as large as possible compared with L_1,
(3) k_1 (the reciprocal of the valve impedance) is as large as possible.

These conclusions are in perfect agreement with the experimental results previously described. In particular it becomes evident at once why the variations of frequency normally observed with Coils II. were much greater than those obtained with Coils I.

Unlimited decrease of C_1 is impossible by reason of the self-capacity of the coil L_1. Again, it is always possible to obtain a valve with a high value of k_1; but this is invariably accompanied by a low amplification factor. To obtain maximum frequency variations, therefore, a valve must be chosen for which $k_1^2 \mu$ is as large as possible.

Unlimited increase of L_2 is impossible by reason of the self-capacity of the latter coil. Thus the natural period

of coil L_2 must never be allowed to become an appreciable fraction of the periodicity of the oscillations, or the theory considered above becomes invalid.

(12) *Conditions for Minimum Frequency Change.*

For minimum frequency change, equation (12), which has been deduced on the assumption that C_1 is very small, becomes invalid. Making use, however, of the condition of maintenance for small frequency changes (equation 11), we may write:

$$\frac{k_1k_3(L_1L_2-M^2)}{L_1C_1}=\frac{k_1^2}{C_1}\left(\frac{L_1L_2}{M^2}-1\right)(\mu M-L_1). \quad (13)$$

$(\mu M-L_1)$, which is always positive in sign, will be made very small by decreasing M and μ and increasing L_1. In these circumstances, $\left(\frac{L_1L_2}{M^2}-1\right)$ will be kept small provided L_2 is made very small. k_1 must be small and C_1 large.

The conditions for minimum frequency variation are therefore exactly the reverse of those given above for maximum variation.

(13) *Non-linearity of Characteristics.*

Up till now we have assumed that the characteristics of the valve could be represented by straight lines. In most practical cases, however, the amplitude of the oscillations is such that the upper or lower bends of the characteristics are reached. It is necessary, therefore, to investigate how this will affect the frequency and the maintenance condition of the oscillations. Taking account of the curvature of the characteristics, we re-write equations (4) and (5) in the forms:

$$i_a=a_0+a(v_a+\mu v_g)+b(v_a+\mu v_g)^2+c(v_a+\mu v_g)^3+\ldots, \quad (4')$$

$$i_g=a_0'+a'(v_a+\mu' v_g)+b'(v_a+\mu' v_g)^2+c'(v_a+\mu' v_g)^3+\ldots. \quad (5')$$

Substituting in these two equations

$$v_a=V\sin pt \quad \text{and} \quad v_g=V'\sin(pt+\eta),$$

and neglecting all frequencies other than "p", we obtain

$$i_a=V\{a+\tfrac{3}{4}c(V^2+2\mu^2V'^2)\}\sin pt + \mu V'\{a+\tfrac{3}{4}c(\mu^2V'^2+2V^2)\}\sin(pt+\eta) \quad (4'')$$

and

$$i_g=V\{a'+\tfrac{3}{4}c'(V^2+2\mu'^2V'^2)\}\sin pt + \mu' V'\{a'+\tfrac{3}{4}c'(\mu'^2V'^2+2V^2)\}\sin(pt+\eta). \quad (5'')$$

Combining these two latter equations with equations (1), (2), and (3), we arrive at the same solution (equations 7 & 8) as was previously obtained, provided we replace k_1, k_2, k_3, and k_4 in the former solution by the expressions:

$$(\alpha) \quad \{a+\tfrac{3}{4}c(V^2+2\mu^2V'^2)+\ldots\},$$

$$(\beta) \quad \mu\{a+\tfrac{3}{4}c(\mu^2V'^2+2V^2)+\ldots\},$$

$$(\gamma) \quad \mu'\{a'+\tfrac{3}{4}c'(\mu'^2V'^2+2V^2)+\ldots\},$$

$$(\delta) \quad \{a'+\tfrac{3}{4}c'(V^2+2\mu'^2V'^2)+\ldots\}, \text{ respectively.}$$

In neglecting all frequencies other than "p", we have assumed that the effect of harmonics on the frequency of the fundamental is negligible, being of the second order.

We see, therefore, that the frequency of the oscillations is dependent upon their amplitude. Physically we may interpret the expressions (α), (β), etc., as representing the "effective" slope of the characteristic over the working range. We shall therefore expect the formulæ previously developed for the frequency of the oscillations to apply, even when working on the bends of the characteristics, provided we replace k_1 etc. by this effective slope. Obvious difficulties would attend accurate measurement of this quantity. For our purpose below, however, the position is simpler, since it is only necessary to have a qualitative knowledge of the effective slope, and to know when it increases or decreases.

Explanation of Experimental Results.

(14) *Variation of Frequency with Filament Current.*

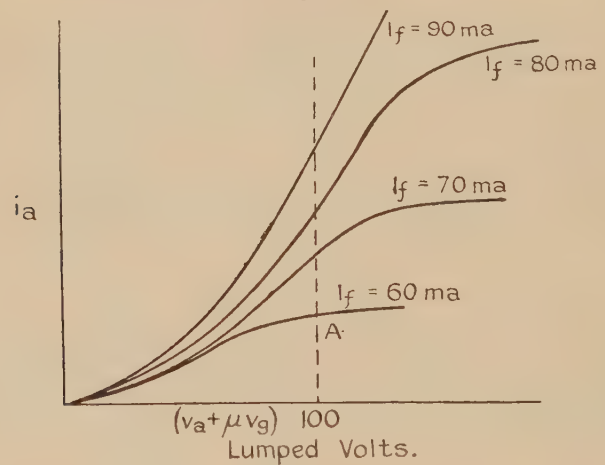
From consideration of equations (8) and (9) it is plain that, provided the values of L_1, L_2, M, and C_1 be kept constant, then the only way in which we can change the frequency of the oscillations is by alteration of the value of k_1. (It is here assumed that μ is always constant, which will be sufficiently correct provided the particular valve used is not changed.) The problem of variation of frequency with filament current reduces, therefore, to the question of how k_1 is affected by variation of filament current. In general, the effect of increase of filament current on the shape of the "lumped"* valve characteristic is to raise the upper or saturation bend, and hence to increase the length of the straight part of the curve. The slope of

* Eccles, Proc. Phys. Soc. xxxii. p. 92 (1920).

the straight part is also slightly increased. A typical set of "lumped" characteristics at various values of filament current are shown in fig. 6.

We may now trace the effect of increase of filament current on the frequency of the oscillations. Imagine the plate potential to be set at 100 volts, and oscillations to commence when the filament current is 60 milliamperes. Then oscillations of small amplitude occur about the representative point A. At this point the slope of the curve (i. e., k_1) is small. Hence the oscillation frequency is not far below

Fig. 6.



the value $\frac{1}{2\pi\sqrt{L_1C_1}}$. As the filament current increases, the representative point passes farther and farther on to the straight part of the curve ; hence k_1 increases and the frequency decreases. At the same time the amplitude of the oscillations increases, and the oscillating point may pass a considerable distance along the volts axis in the negative direction. Eventually a stage is reached at which the saturation bend becomes so high that it is not reached by the oscillating point.

If we commence with a higher plate voltage, then we need to make a correspondingly larger increase in filament current before getting on to the straight part of the characteristic. Hence with high plate voltages the initial rate of decrease of frequency is slow.

At high plate voltages the negative excursion of the oscillating point is less. Hence the effective slope of

the characteristic is greater, and the total frequency change greater at high plate voltages.

When the filament current becomes so great that the oscillating point ceases to reach the upper bend of the characteristic, the rate of decrease of frequency again falls off.

The Frequency Minimum.

For the production of a minimum frequency on the frequency-filament current curve the chief necessary condition is that the oscillating point shall not reach the saturation bend of the characteristic. For if the bend is reached, then increase of filament current, by raising the position of the bend, will increase the effective slope k_1, and so decrease the frequency continuously.

There is experimental evidence to show that, for high values of filament current, the effect of increase of filament current is to flatten out the bottom bend of the characteristic. This is probably accounted for by the strong "space-charge" present. In these circumstances, owing to the flattening of the bottom bend, the effect of increase of filament current is to reduce the effective slope of the characteristic, and consequently to raise the frequency of the oscillations. For the same reason the amplitude of the oscillations is reduced simultaneously.

It is now clear why loose coupling, low plate voltage, and high filament current are favourable conditions for the production of a minimum frequency. All of these conditions tend to ensure that the oscillating point shall not reach the upper bend. At the same time, a high filament current is necessary in order to produce a large emission of electrons, with the consequent flattening of the bottom bend. Further, it is now clear why a valve, after working for some time at high filament current, may fail to give rise to a minimum frequency. The filament has "aged," and the emission is no longer sufficient to give rise to the effect.

(15) *Variation of Frequency with Plate Voltage.*

In this case also the variations of frequency may all be traced to variation of k_1. Let us imagine (fig. 6) that oscillations are started at a low plate voltage. Here k_1 has a low value, and the oscillations have very nearly the frequency $\frac{1}{2\pi\sqrt{L_1C_1}}$. Increase of plate voltage removes the representative point away from the bottom bend,

so that k_1 increases and the frequency falls. Eventually, as the plate voltage is increased, the representative point approaches the upper bend, and the frequency commences to rise. If the amplitude of the oscillations is such that the oscillating point can reach both bends at once, then the minimum frequency occurs at a value of plate voltage about midway between the bends. Hence increase of filament current raises the position of this minimum frequency.

If the amplitude of the oscillations is small, then the oscillating point will not normally reach the upper bend of the characteristic. Hence increase of plate voltage will produce a continuous fall in frequency until a high voltage is reached sufficiently near to the upper bend. Hence the minimum frequency, for oscillations of small amplitude, will occur at a higher value of plate voltage than that for large oscillations. This explains why reduction of the coupling has the effect of shifting the position of the frequency minimum to a higher value of plate voltage. For higher values of filament current the length of the straight part of the characteristic is greater. Hence the effective slope k_1 is greater and the total frequency lowering greater.

(16) *Variation of Frequency with Coupling.*

The theoretical effect of coupling on the frequency is given by equation (13). For small frequency variations, by differentiation we get

$$\frac{dp}{d\mathrm{M}} = \frac{k_1^2 \mathrm{L}_1 p^3}{2\mathrm{M}^3} \{\mu \mathrm{M}^3 + \mu \mathrm{L}_1 \mathrm{L}_2 \mathrm{M} - 2\mathrm{L}_1^2 \mathrm{L}_2\}. \quad . \quad (14)$$

Taking the numerical values for L_1 etc., appropriate for Coils I., and plotting the function for different values of M, we obtain the following results. For low values of M, $\frac{dp}{d\mathrm{M}}$ is negative. It is zero at a value of M about the middle of the scale, and positive for high values of M. This corresponds to the experimental curve type (*b*) (fig. 4). This type therefore, obtained at medium values of plate voltage, is the type of curve normally to be expected on theoretical grounds.

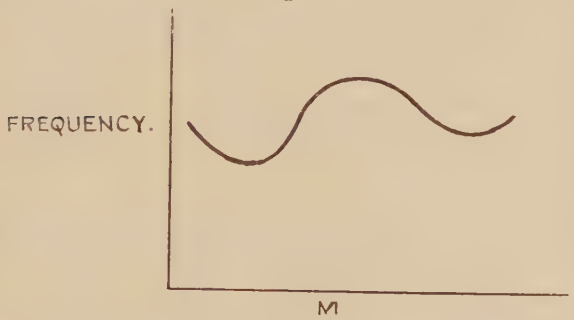
In order to explain types (*a*) and (*c*), we require to consider the amplitude of the oscillations. It is well known that, if M be steadily increased, then the amplitude of the oscillations first increases from zero, reaches a maximum value, and thereafter decreases. If the representative point is near to one of the bends on the characteristic,

then increase of amplitude sends the oscillating point round this bend, so reduces the value of the effective slope k_1, and increases the frequency. Hence if the plate voltage is such that the representative point is near to a bend in the characteristic, then the frequency will vary in exactly the same manner as the amplitude.

We see, therefore, that when we alter the value of M we have two influences operating to produce a change of frequency.

One influence, that due to equation (14), tends to produce a minimum frequency at about the middle of the curve, while the other, that due to change of amplitude, tends to produce a maximum frequency at the same position. The resultant of these two effects will produce a frequency variation of the type shown in fig. 7. This curve was

Fig. 7.



obtained experimentally in only one case, while using a high plate voltage (type *c'* above). Adjustment of the plate voltage to a value approximately midway between the bends of the characteristic tends to eliminate the effect due to change of amplitude (type *b*).

Type (*a*) is produced by proximity to the bottom bend of the characteristic. The amplitude at low values of M is in this case insufficient to produce an initial minimum.

Type (*c*) is that usually produced by proximity to the upper or saturation bend of the characteristic.

(17) *Possible Applications.*

Several applications of the effects described above suggest themselves. Thus it is possible to amplify small changes of current into very large changes of frequency. This effect may be utilized in order to measure small high-frequency currents. The current to be measured is applied to the filament of the valve, either through a transformer,

or directly, with choke coils in series with the filament battery. Since it is possible to produce a change of frequency of several octaves by an alteration of only a few milliamperes in filament current, the method appears capable of great sensitivity. It possesses the advantages of giving a reading independent of frequency, and of introducing a negligibly small impedance into the circuit carrying the current to be measured. It remains to be seen if the filament emission remains sufficiently constant to ensure reasonable permanence of calibration.

The writer has not yet had the opportunity to make a thorough examination of the possibilities of the method, but intends doing so as soon as circumstances permit.

Summary.

Frequency variations due to changes in filament current, plate voltage, and coupling are investigated experimentally. Variations as great as two octaves or more are described. The most important cause of frequency variations is found to be the flow of grid current. The conditions for maximum and for minimum frequency change are worked out. The experimental results are then explained in detail by means of the theory.

In conclusion, I wish to express my sincere thanks to Professor E. Taylor Jones for his very kind interest and advice, and to the University of Glasgow for the facilities to carry out the above research.

Natural Philosophy Department,
University of Glasgow.
July 1927.

[*Addendum.*—While this paper was going through the press the attention of the author was drawn to a paper by Strecker (*Jahrb. Draht. Teleg.* xxii. (1923)) on the same subject. This valuable paper had escaped the notice of the present writer since no reference to it had appeared in 'Science Abstracts.' In this paper Strecker considers the effect of grid current, and arrives at equations equivalent to nos. 6, 7, and 8 above. In his work Strecker has confined his attention to oscillations of very small amplitude, and with the results which he obtains in these circumstances the present writer is in entire agreement. In the present paper, however, oscillations of large amplitude are considered, and for this reason the method of treatment and the nature of the experimental results differ considerably from those of Strecker.]

LXXXVIII. *On Maxwell's Stress, and its Time Rate of Variation.* *By* S. R. Milner, *D.Sc., F.R.S., Professor of Physics, The University, Sheffield* *.

IN the electromagnetic field three functions of **e** and **h** are known to which dynamical meanings can be ascribed. They are †

$$W=\tfrac{1}{2}(e^2+h^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

the energy density,

$$\mathbf{G}=\frac{1}{c}[\mathbf{eh}], \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

the momentum density, and Maxwell's stress system defined by

$$\Pi_n=\tfrac{1}{2}(e^2+h^2)\mathbf{n}-(\mathbf{en})\mathbf{e}-(\mathbf{hn})\mathbf{h}. \quad . \quad . \quad . \quad (3)$$

Π_n stands for the stress exerted across any unit plane whose normal is **n**, positive values indicating pressures. The justification for calling this function of **e**, **h**, and **n** a stress system is found in the following relation between Π_n and **G**, which can be derived from the fundamental equations:

$$\frac{\partial G_x}{\partial t}=-\left(\frac{\partial \Pi_{xx}}{\partial x}+\frac{\partial \Pi_{yx}}{\partial y}+\frac{\partial \Pi_{zx}}{\partial z}\right), \text{ etc.} \quad . \quad . \quad (4)$$

This shows that the time rate of increase of the electromagnetic momentum in a fixed unit volume is equal to the resultant force which would be exerted by Π_n on that volume if Π_n were an ordinary mechanical stress.

Maxwell, I believe, did not doubt the existence of a mechanical stress actually operating on the medium, or, as we should now say preferably, on the field. Later writers have questioned its actuality. The fact is that Π_n satisfies one condition which we postulate as a property of a mechanical stress, but it does not satisfy another and an equally necessary one. Π_n is by definition a quantity at a *fixed point* in the observer's space; it varies with the time, but it is not conceived as a quantity which *moves*. Consequently, although it accounts in accordance with dynamical principles for the variation of the momentum of the field, it is incapable of doing work and accounting for the time variation of the energy ‡. An expression for the rate of increase of the

* Communicated by the Author.

† In this paper Heaviside units are used; [..] denotes the vector product; scalar products are not marked with special signs, round brackets and dots being used with the ordinary meanings.

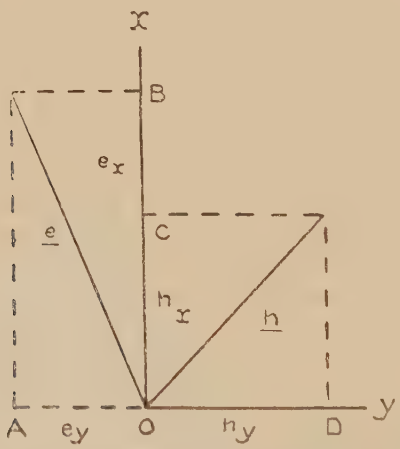
‡ To do this a *moving stress* is required, and this differs in its magnitude from Maxwell's Π_n. *Cf.* E. Cunningham, Proc. Roy. Soc. A, lxxxiii. p. 110 (1910); S. R. Milner, *loc. cit.* cxiv. p. 23 (1927).

energy in a fixed unit volume can indeed be found, but it comes out, not in terms of Π_n, but of $\mathbf{G}$, viz. Poynting's theorem :

$$\frac{\partial W}{\partial t} = -c^2\left(\frac{\partial G_x}{\partial x} + \frac{\partial G_y}{\partial y} + \frac{\partial G_z}{\partial z}\right). \quad . \quad . \quad . \quad (5)$$

In spite of the fact that it does not satisfy fully the dynamical requirements of a stress, Π_n is a useful and important function of $\mathbf{e}$ and $\mathbf{h}$, which can be imagined as operating at each point of an electromagnetic field and producing the changes which take place there. It is difficult to visualize Π_n in its vectorial form (3) ; it is easy, however, if we consider it as a system of principal stresses, each of which has a definite magnitude and acts along one of three rectangular axes suitably oriented at each point of the field.

Fig. 1.



Consider any point of the field and take the plane containing $\mathbf{e}$ and $\mathbf{h}$ as that of xy. Draw axes Ox, Oy in this plane such that the rectangles AB, CD (fig. 1) are equal ; *i. e.*, such that

$$e_x e_y + h_x h_y = 0.$$

Then Ox, Oy, and the perpendicular Oz are the principal axes of Maxwell's stress at the point. Only when $\mathbf{n}$ is along one of these directions is the stress normal to the area on which it acts. Π_1, Π_2, Π_3 being the principal stresses along Ox, Oy, Oz, it may easily be shown that

$$\left.\begin{aligned} -\Pi_1 = +\Pi_2 &= \tfrac{1}{2}\{(e^2+h^2)^2 - 4[\mathbf{eh}]^2\}^{\frac{1}{2}} \\ &= (W^2 - c^2G^2)^{\frac{1}{2}}, \\ \Pi_3 &= \tfrac{1}{2}(e^2+h^2) = W. \end{aligned}\right\} \quad . \quad . \quad (6)$$

Π_1 is a tension, Π_2 and Π_3 are pressures, Π_2 being numerically equal to Π_1. The figure formed by three perpendicular lines, suitably oriented at each point of the field, together with the magnitude of the principal stresses along them, forms a complete geometrical representation of a three-dimensional symmetrical tensor of the second rank in the same way that a line of given magnitude and orientation represents an ordinary vector.

Instead of the principal stresses, it is more convenient to use two variables simply related to them, which we may call "stress vectors," defined thus:

$$\left.\begin{aligned} p &= (\Pi_3 - \Pi_1)^{\frac{1}{2}} \text{ acting along } Ox, \\ q &= (\Pi_3 - \Pi_2)^{\frac{1}{2}} \quad \text{,,} \quad \text{,,} \quad Oy. \end{aligned}\right\} \quad . \quad . \quad . \quad (7)$$

p and **q** are vectors of the same type as **e** and **h**, which are by definition at right angles to each other, whether **e** and **h** are or are not. They specify completely the dynamical variables by

$$\left.\begin{aligned} -\Pi_1 = +\Pi_2 &= \tfrac{1}{2}(p^2 - q^2), \\ \Pi_3 = \mathrm{W} &= \tfrac{1}{2}(p^2 + q^2), \\ \mathbf{G} &= \frac{1}{c}[\mathbf{pq}], \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (8)$$

and are also related to **e** and **h** by the equations

$$\left.\begin{aligned} \mathbf{p} &= \quad \mathbf{e}\cos\alpha + \mathbf{h}\sin\alpha, \\ \mathbf{q} &= -\mathbf{e}\sin\alpha + \mathbf{h}\cos\alpha, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where

$$\tan 2\alpha = \frac{2(\mathbf{eh})}{e^2 - h^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The fact that the time rates of increase of W and **G** can be expressed in terms of the space variations of **G** and Π_n (eqs. (5) and (4)) suggests a search for an expression for the time rate of variation of Π_n. If this could be obtained in terms of W, **G**, and Π_n, the three equations would form a closed chain, giving a complete formulation of the electromagnetic laws in terms of dynamically interpretable quantities. It appears, however, that in the most general field (where **e** and **h** are not perpendicular) this cannot be done until another variable has been introduced into the equations. W, **G**, and Π_n together specify only 5 independent variables between them—*e. g.*, the magnitudes of p and q, together with 3 direction cosines to fix the orientation of the rectangular stress axes. The complete field, on the other hand, at each point requires 6 independent variables

to express it—*e. g.*, the magnitudes of e and h, with 4 direction cosines to fix their lines of action. Introducing the α of equation (10) as a sixth variable, we see that $\mathbf{p}$, $\mathbf{q}$, and α will fully specify all quantities, both dynamical and electromagnetic (except electric charges), in the field.

The time variation of Maxwell's stress appears to be most simply expressed by formulæ stating separately the time rates of increase of p^2 and q^2 (which fix the magnitudes of the principal stresses), and the angular velocity with which the system of stress axes is rotating *.

On substituting from (9) in the fundamental equations

$$\dot{\mathbf{e}} = c \operatorname{curl} \mathbf{h}, \quad \dot{\mathbf{h}} = -c \operatorname{curl} \mathbf{e},$$

we obtain equivalent ones in terms of p, q, and α:

$$\left.\begin{aligned} \dot{\mathbf{p}} - \mathbf{q}\dot{\alpha} &= c(\operatorname{curl} \mathbf{q} - [\mathbf{p}\nabla\alpha]), \quad (a) \\ \dot{\mathbf{q}} + \mathbf{p}\dot{\alpha} &= -c(\operatorname{curl} \mathbf{p} + [\mathbf{q}\nabla\alpha]). \quad (b) \end{aligned}\right\} \quad (11)$$

Since $\mathbf{p}$ and $\mathbf{q}$ are everywhere perpendicular, α can be eliminated by multiplying scalarly by $\mathbf{p}$ and $\mathbf{q}$, giving

$$\left.\begin{aligned} \frac{\partial}{\partial t}(\tfrac{1}{2}p^2) &= c\mathbf{p} \operatorname{curl} \mathbf{q}, \\ \frac{\partial}{\partial t}(\tfrac{1}{2}q^2) &= -c\mathbf{q} \operatorname{curl} \mathbf{p}. \end{aligned}\right\} \quad (12)$$

These are just the same relations as exist between $\mathbf{e}$ and $\mathbf{h}$, although $\mathbf{p}$ and $\mathbf{q}$ are vectors quite different from $\mathbf{e}$ and $\mathbf{h}$ in both magnitude and direction.

Although α does not appear in (12), it cannot be eliminated from expressions for the rotation of the stress axes. A complete set of equations can, however, be derived, three to give at a fixed point the component velocities of angular rotation of the stress system about its own axes in terms of the space derivatives of $\mathbf{p}$, $\mathbf{q}$, and α, and a fourth the time rate of increase of α in terms of space derivatives of $\mathbf{p}$ and $\mathbf{q}$.

The equation for $\dot{\alpha}$ may readily be obtained by multiplying (11 a) by $\mathbf{q}$ and (11 b) by $\mathbf{p}$ and adding. Since

$$\mathbf{q}[\mathbf{p}\nabla\alpha] + \mathbf{p}[\mathbf{q}\nabla\alpha] = 0 \quad \text{and} \quad \mathbf{q}\dot{\mathbf{p}} + \mathbf{p}\dot{\mathbf{q}} = 0,$$

we obtain at once

$$\frac{\partial \alpha}{\partial t} = c\,\frac{(\mathbf{q} \operatorname{curl} \mathbf{q} - \mathbf{p} \operatorname{curl} \mathbf{p})}{p^2 - q^2}. \quad (13)$$

To obtain the rotations, use for the moment the special

* The attempt to express $\frac{\partial \Pi_n}{\partial t}$ direct leads to great complications.

axes Ox, Oy, Oz of fig. 1, which are along the directions of $\mathbf{p}$, $\mathbf{q}$ and $[\mathbf{pq}]$ respectively. On these axes

$$p_x=p, \quad p_y=p_z=0, \quad q_y=q, \quad q_x=q_z=0.$$

If we write $d\theta_{yz}$, $d\theta_{zx}$, $d\theta_{xy}$ for the infinitesimal angles through which the directions of $\mathbf{p}$, $\mathbf{q}$ and $[\mathbf{pq}]$ have rotated at a neighbouring point of space or time, where the components of $\mathbf{p}$ and $\mathbf{q}$ are p_x+dp_x, etc., we have by simple geometry

$$\begin{aligned} &dp_x=dp, && dp_y=p\,d\theta_{xy}, && dp_z=-p\,d\theta_{zx}, \\ &dq_x=-q\,d\theta_{xy}, && pq_y=dq, && dq_z=q\,d\theta_{yz}. \end{aligned}$$

We can consequently write for the component angular velocities of the stress system

$$\omega_x=\frac{\partial\theta_{yz}}{\partial t}=\frac{1}{q}\frac{\partial q_z}{\partial t},$$

$$\omega_y=\frac{\partial\theta_{zx}}{\partial t}=-\frac{1}{p}\frac{\partial p_z}{\partial t},$$

$$\omega_z=\frac{\partial\theta_{xy}}{\partial t}=\frac{1}{p}\frac{\partial p_y}{\partial t}=-\frac{1}{q}\frac{\partial q_x}{\partial t}.$$

Since $\mathbf{p}$, $\mathbf{q}$, and $[\mathbf{pq}]$ are perpendicular, these equations can be written in vectorial form independent of the special axes:

$$\left.\begin{aligned} q^2\,.\,\mathbf{p}\,\omega &= [\mathbf{pq}]\dot{\mathbf{q}}, \quad . \quad . \quad (a) \\ p^2\,.\,\mathbf{q}\,\omega &= -[\mathbf{pq}]\dot{\mathbf{p}}, \quad . \quad . \quad (b) \\ [\mathbf{pq}]\,\omega &= \mathbf{q}\dot{\mathbf{p}}=-\mathbf{p}\dot{\mathbf{q}}. \quad . \quad (c) \end{aligned}\right\} \quad . \quad . \quad . \quad (14)$$

ω stands here for the vector angular velocity of the system of stress axes. (14 a) gives, on substituting from (11 b) for $\dot{\mathbf{q}}$ *,

$$q^2\,.\,\mathbf{p}\,\omega=c(-[\mathbf{pq}]\,\mathrm{curl}\,\mathbf{p}+q^2\,.\,\mathbf{p}\nabla\alpha). \quad . \quad (15\,a)$$

Similarly, from (14 b),

$$p^2\,.\,\mathbf{q}\,\omega=c(-[\mathbf{pq}]\,\mathrm{curl}\,\mathbf{q}+p^2\,.\,\mathbf{q}\nabla\alpha), \quad . \quad (15\,b)$$

and from (14 c)

$$\begin{aligned}[\mathbf{pq}]\,\omega &= c(\mathbf{q}\,\mathrm{curl}\,\mathbf{q}-\mathbf{q}[\mathbf{p}\nabla\alpha])+q^2\dot{\alpha} \\ &= c\left(\frac{p^2\,.\,\mathbf{q}\,\mathrm{curl}\,\mathbf{q}-q^2\,.\,\mathbf{p}\,\mathrm{curl}\,\mathbf{p}}{p^2-q^2}+[\mathbf{pq}]\nabla\alpha\right). \quad (15\,c)\end{aligned}$$

on substituting for $\dot{\alpha}$ from (13).

(15 a, b, c) determine the three scalar products of ω with $\mathbf{p}$, $\mathbf{q}$, and $[\mathbf{pq}]$ respectively—*i. e.*, the component angular

* Since $[\mathbf{pq}][\mathbf{q}\nabla\alpha]=-\nabla\alpha[\mathbf{q}[\mathbf{pq}]]$ and $[\mathbf{q}[\mathbf{pq}]]=q^2\,.\,\mathbf{p}$.

velocity of the stress system about each of its own axes is derived in terms of the space variations of the stress vectors and of α. The complete angular velocity is thus determined.

The point of chief interest about these equations is the remarkable way in which $\nabla\alpha$ enters into them. They may be written in the form

$$\left.\begin{aligned} \mathbf{p}(\omega - c\nabla\alpha) &= -c\frac{[\mathbf{pq}]\operatorname{curl}\mathbf{p}}{q^2}, \\ \mathbf{q}(\omega - c\nabla\alpha) &= -c\frac{[\mathbf{pq}]\operatorname{curl}\mathbf{q}}{p^2}, \\ [\mathbf{pq}](\omega - c\nabla\alpha) &= c\frac{p^2\cdot\mathbf{q}\operatorname{curl}\mathbf{q} - q^2\cdot\mathbf{p}\operatorname{curl}\mathbf{p}}{p^2 - q^2}. \end{aligned}\right\} \quad . \quad . \quad (16)$$

or still more simply, denoting by suffixes components on the special axes defined by the stress,

$$\left.\begin{aligned} (\omega - c\nabla\alpha)_x &= -\frac{c}{q}\operatorname{curl}_z\mathbf{p}, \\ (\omega - c\nabla\alpha)_y &= -\frac{c}{p}\operatorname{curl}_z\mathbf{q}, \\ (\omega - c\nabla\alpha)_z &= c\frac{p\operatorname{curl}_y\mathbf{q} - q\operatorname{curl}_x\mathbf{p}}{p^2 - q^2}. \end{aligned}\right\} \quad . \quad . \quad . \quad (17)$$

Write $\omega' = \omega - c\nabla\alpha$, so that $\omega = \omega' + c\nabla\alpha$. ω' is expressed solely in terms of $\mathbf{p}$, $\mathbf{q}$, and their space derivatives, *i. e.*, is completely defined by the stress. In a field in which $\alpha = 0$ everywhere (a restriction which implies that $\mathbf{e}$ and $\mathbf{h}$ shall be everywhere perpendicular to each other), ω' is the full rotation which the stress system possesses. In the more general field, however, in which $\mathbf{e}$ and $\mathbf{h}$ are not perpendicular, to obtain the actual angular velocity of the stress axes, there has to be added to ω' an additional rotation, $c\nabla\alpha$. α is a scalar variable which, as has been pointed out, is entirely independent of the stress. It is simply related to $(\mathbf{eh})$, and from the electromagnetic point of view forms a measure at each point of the extent to which the field deviates from the condition that $\mathbf{e}$ and $\mathbf{h}$ are perpendicular. We see now that it has also a certain dynamical significance; α is a potential whose gradient (multiplied by c) constitutes a velocity of rotation of the stress axes which exists in the general field independently of and superposed on that (ω') produced by the stress itself. The presence of this "intrinsic" rotation, as it may be called, sharply distinguishes the properties of fields in which

e and **h** are not perpendicular and fields in which they are. The matter may be put in this way. Consider any "orthogonal" field (**h**⊥**e** everywhere). A knowledge of the stress system at each point enables its rate of increase (in free space) to be determined completely, and with it the whole subsequent changes in the field. In a non-orthogonal field this is not the case. There is now present at each point an intrinsic rotation of the stress system depending for its amount on the extent to which the field deviates from orthogonality, and the changes in the field can only be fully predicted from a knowledge of the stress system when the effect of this rotation is taken into account.

LXXXIX. *The Relativistic Rule for the Equipartition of Energy.* *By* F. F. P. Bisacre, *O.B.E., M.A.**

THE well-known rule of equipartition of energy in a gas is that each degree of freedom carries energy, $\frac{1}{2}$RT.

If classical mechanics be abandoned in favour of the mechanics of the special theory of relativity, this rule must be modified for high temperatures. It is of some interest, perhaps, to examine the nature of the modification needed. The modification turns out to be a very simple one that is accurately expressible in terms of Bessel Functions of the second kind.

Consider a mass of gas, consisting of n particles, each of proper mass, m, and being at temperature T.

Assume that the potential energy is negligible so that the whole energy (*i. e.*, the kinetic energy and the "internal" energy) of the system is given by

$$\mathrm{E} = mc^2 \sum_i \left\{ 1 - \frac{\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2}{c^2} \right\}^{-\frac{1}{2}}, \quad i = 1, 2, \ldots n \quad . \quad (1)$$

in the usual notation.

The kinetic energy, κ, is given by

$$\kappa = \mathrm{E} - mnc^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Plot a 3-dimensional velocity diagram having one representative point per particle. The coordinates of a point give the component speeds of the particle it stands for.

The radius vector, r, gives the magnitude of the resultant velocity of the particle; hence $r < c$: in fact, the whole velocity diagram lies inside a sphere of radius c.

* Communicated by the Author.

Let ϵ be the whole energy of a single particle, then

$$\epsilon = mc^2 \left\{ 1 - \frac{r^2}{c^2} \right\}^{-\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

so that the representative points of all particles having energy ϵ must lie on the spherical surface of radius

$$r = c \left\{ 1 - \frac{m^2 c^4}{\epsilon^2} \right\}^{\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and consequently all particles having energy between ϵ and $\epsilon + \delta\epsilon$ have representative points lying between the spherical surfaces of radii r and $r + \delta r$, where

$$\delta r = m^2 c^5 \epsilon^{-3} \left\{ 1 - \frac{m^2 c^4}{\epsilon^2} \right\}^{-\frac{1}{2}} \delta\epsilon. \quad . \quad . \quad . \quad . \quad (5)$$

Since the velocity cannot exceed c, we must carefully reconsider the usual conception of "equal probability." I follow Cunningham * in this :—

"If nothing is known of a given particle save that it is in a certain region of space, it is assumed that all positions within that region are equally likely to be the actual position of the particle ; or, again, it is assumed that for a given particle of which nothing is known to restrict its velocity, all velocities are equally probable, no matter how great. Now, in the case of the velocity it is obvious that the principle of relativity cannot admit this as a reasonable assumption, since the continual addition of velocities never leads to a velocity greater than that of light ; and so the question may be asked : 'What criteria of equal probability are consistent with the principle?'"

In classical theory, if $\delta\tau$ is an element of volume of the velocity diagram corresponding to u, $u + \delta u$; v, $v + \delta v$; w, $w + \delta w$, and we find n representative points in this volume, then we must suppose that, if nothing is known to restrict the velocity, we shall find the same number of points in another volume the same size at some other place in the diagram, u', v', w'.

This rule cannot hold where the velocity cannot be greater than c.

"A general criterion may be laid down applying to all cases.

> "Any two states of a self-contained system which can be transformed into one another by a Lorentz transformation are to be considered as equally probable."

By comparing a particle with a small velocity (u, v, w)

* E. Cunningham, 'The Principle of Relativity,' p. 208 (1921).

with one obtained from it by a Lorentz transformation, it follows that, other things being equal, it is just as probable that the velocity of any particular particle will be represented by a point within the volume $\delta\tau$, near the origin, as within $\delta\tau'$ elsewhere, where

$$\delta\tau' = \delta\tau \,.\, \beta^{-4}$$

and $$\beta = \left[1 - \frac{v^2}{c^2}\right]^{-\frac{1}{2}} = \epsilon/mc^2 \text{ by (3)},$$

where v is the resultant velocity, *i. e.* the radius, r.

Dividing up the space enclosed by the spherical surface $r=c$ into cells of equal probability and distributing the points in the most probable way *so as to keep the total number constant and the total energy they represent constant*, we find, in the usual way *, that the number of representative points lying in the cell of volume $\tau\beta^{-4}$ is $Ae^{-\gamma\epsilon}$, where A and γ are constants and ϵ is the energy corresponding to a point in the cell.

The number of such cells in the spherical shell of radii r and $r+\delta r$ is

$$\frac{4\pi r^2 \,.\, \delta r \,.\, \beta^4}{\tau}.$$

If δn is the number of representative points lying within this shell,

$$\delta n = \frac{4\pi r^2 \,.\, \delta r \,.\, \beta^4}{\tau} \,.\, Ae^{-\gamma\epsilon} = Br^2\beta^4 e^{-\gamma\epsilon}\delta r,$$

where B is a constant, and δn is the number of particles having energy lying between ϵ and $\epsilon+\delta\epsilon$.

Substituting known values for β, r, and δr, we get

$$\delta n = \frac{B}{m^2c}\left\{1 - \frac{m^2c^4}{\epsilon^2}\right\}^{\frac{1}{2}} \epsilon e^{-\gamma\epsilon}\delta\epsilon\,;$$

i. e., $$n = \frac{B}{m^2c}\int_{mc^2}^{\infty}\left\{1 - \frac{m^2c^4}{\epsilon^2}\right\}^{\frac{1}{2}} \epsilon e^{-\gamma\epsilon}d\epsilon = \frac{B}{m^2c} \,.\, I_d$$

and $$E = \frac{B}{m^2c}\int_{mc^2}^{\infty}\left\{1 - \frac{m^2c^4}{\epsilon^2}\right\}^{\frac{1}{2}} \epsilon\, e^{-\gamma\epsilon}d\epsilon = \frac{B}{m^2c} \,.\, I_n.$$

$\therefore$ $$E = nI_n/I_d.$$

The ordinary kinetic energy, κ, is given by

$$\kappa = n\left\{\frac{I_n}{I_d} - mc^2\right\}.$$

There are $3n$ degrees of freedom ; hence, if ϵ is the quota,

$$\epsilon = \frac{1}{3}\left\{\frac{I_n}{I_d} - mc^2\right\}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

* J. H. Jeans, 'The Dynamical Theory of Gases,' p. 45 *et seq.* (1921).

It remains to evaluate the integrals.

The substitutions

$$\epsilon = mc^2 x \quad \text{and} \quad \alpha = \gamma mc^2$$

lead to

$$I_n/I_d = mc^2 \,.\, F(\alpha)/\psi(\alpha),$$

where

$$F(\alpha) = \int_1^\infty \left\{1 - \frac{1}{x^2}\right\}^{\frac{1}{2}} x^2 e^{-\alpha x}\, dx$$

and

$$\psi(\alpha) = \int_1^\infty \left\{1 - \frac{1}{x^2}\right\}^{\frac{1}{2}} x e^{-\alpha x}\, dx.$$

Now *

$$F(\alpha) = \left[\int_1^\infty (x^2-1)^{-\frac{1}{2}} x^3 e^{-\alpha x}\, dx - \int_1^\infty (x^2-1)^{-\frac{1}{2}} x e^{-\alpha x}\, dx\right]$$

and

$$\psi(\alpha) = \left[\int_1^\infty (x^2-1)^{-\frac{1}{2}} x^2 e^{-\alpha x}\, dx - \int_1^\infty (x^2-1)^{-\frac{1}{2}} e^{-\alpha x}\, dx\right],$$

and each of these integrals can be obtained from

$$f(\alpha) = \int_1^\infty (x^2-1)^{-\frac{1}{2}} e^{-\alpha x}\, dx$$

by differentiation; in fact

$$F(\alpha) = [f'(\alpha) - f'''(\alpha)]$$

and

$$\psi(\alpha) = [f''(\alpha) - f(\alpha)],$$

where accents denote differentiation with respect to α.

By putting $\cosh\theta$ for x we can reduce $f(\alpha)$ to

$$f(\alpha) = \int_0^\infty e^{-\alpha\cosh\theta}\, d\theta, \quad \ldots\ldots \quad (7)$$

which is the Bessel function of the second kind, defined by

$$f(\alpha) = K_0(\alpha) = (\log_e 2 - 0{\cdot}577\ldots)\, J_0(i\alpha) - J_0(i\alpha)\log_e \alpha + \left[\frac{\alpha^2}{2^2} + \ldots\right],$$

and is Heine's function. It satisfies the related Bessel equation:

$$\frac{d^2w}{d\alpha^2} + \frac{1}{\alpha}\frac{dw}{d\alpha} - w = 0. \quad \ldots\ldots \quad (8)$$

* I am indebted to Dr. John Dougall for the suggestion to split up the integrals in this way.

The function is tabulated on pp. 127, 128 of Jahnke and Emde's 'Funktionentafeln'—so is the function $K_1(\alpha)$, which is the same as $-\frac{dK_0}{d\alpha}$, *i.e.* as $-f'(\alpha)$.

We have, then,

$$F(\alpha)/\psi(\alpha)=\left[\frac{K_0(\alpha)}{K_1(\alpha)}+\frac{2}{\alpha}\right] \quad . \quad . \quad . \quad . \quad (8a)$$

by using equation (8) to express $K_0''(\alpha)$ and $K_0'''(\alpha)$, in terms of $K_0(\alpha)$ and $K_0'(\alpha)$, *i.e.* $-K_1(\alpha)$.

The result is that

$$\bar{\epsilon}=\frac{mc^2}{3}\left[\frac{K_0(\alpha)}{K_1(\alpha)}+\frac{2}{\alpha}-1\right], \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where $\alpha=\gamma mc^2$.

The asymptotic value of $K_0(\alpha)$ is given by

$$K_0(\alpha)\approx\sqrt{\frac{\pi}{2}}\cdot\frac{e^{-\alpha}}{\sqrt{\alpha}}\left\{1-\frac{1}{8\alpha}+\ldots\right\}$$

from which it appears that

$$\bar{\epsilon}=\frac{1}{2\gamma} \text{ when } \alpha \text{ is large.}$$

Now c is large when T is small, so that this is the classical result, and we know that, *when* T *is small*, $\bar{\epsilon}=\frac{1}{2}RT$, so that γ is $1/RT$, which formula we shall assume to hold good for all values of T.

$$\bar{\epsilon}=\tfrac{1}{2}RT,$$

$$\therefore \quad \gamma=1/RT.$$

The general result is, then,

$$\left.\begin{aligned}\bar{\epsilon}&=\frac{mc^2}{3}\left[\frac{K_0(\alpha)}{K_1(\alpha)}+\frac{2}{\alpha}-1\right],\\ \text{where}\quad \alpha&=\frac{mc^2}{RT}.\end{aligned}\right\} \quad . \quad . \quad . \quad (10)$$

For a given gas and temperature, α is calculable, and values of $K_0(\alpha)$ and $K_1(\alpha)$ are given in Jahnke and Emde's Tables.

When α is small, the important term in the Bessel function is the logarithm of α, and we get

$$K_0(\alpha)\approx-\log\alpha$$

and $$K_1(\alpha)\approx 1/\alpha.$$

$$\therefore \quad \bar{\epsilon}\rightarrow 2/3\gamma,$$

since $\alpha\log\alpha\rightarrow 0$ as $\alpha\rightarrow 0$.

We thus have

$$\left.\begin{aligned} \bar{\epsilon} &\to \tfrac{2}{3}\mathrm{RT} \text{ as } \mathrm{T}\to\infty\,, \\ \bar{\epsilon} &\to \tfrac{1}{2}\mathrm{RT} \text{ as } \mathrm{T}\to 0. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (11)$$

Conclusion.

It is sometimes stated that the classical rule, $\frac{1}{2}\mathrm{RT}$, is "inevitable," whatever system of dynamics be used, so long as continuity is preserved—to break away from the classical rule discontinuities are *essential*. This is plainly not true. In breaking away from classical *kinematics*, the criterion of equal probability comes in question, so that the simple exponential rule is no longer applicable. If one tries to apply the simple exponential rule to this problem, one reaches the absurdity that the energy of the gas is *finite* as $\mathrm{T}\to\infty$. Cunningham's modified probability rule avoids this, and these results go to show the essential reasonableness of his rule.

Even in a field free from gravitational forces, the classical rule is merely an approximation (though, of course, a very good one) to a much more complicated law. The specific heat, even of an *ideal* gas, must show an upward tendency (of the amount indicated by the above results) which arises purely from the limitations imposed by the nature of space-time.

March 1, 1927.

Note added in proof.—It may be held that the proper probability rule to take involves not β^4 but β^5 (see Cunningham, 'The Principle of Relativity,' 1921, p. 209, §5).

If this rule be taken,

$$\frac{1}{mc^2}\frac{\mathrm{E}}{n} = \frac{1}{mc^2}\frac{\mathrm{I}_n}{\mathrm{I}_d} = \frac{(\alpha^2+6)\mathrm{K}_1(\alpha)+3\alpha\mathrm{K}_0(\alpha)}{2\alpha\mathrm{K}_1(\alpha)+\alpha^2\mathrm{K}_0(\alpha)}$$

giving

$$\bar{\epsilon} = 1/2\gamma \text{ when } \alpha \text{ is large}$$

and $$\bar{\epsilon} = 1/\gamma \text{ when } \alpha \text{ is small};$$

i. e.,

$$\bar{\epsilon} \longrightarrow \tfrac{1}{2}\mathrm{RT} \text{ as } \mathrm{T}\longrightarrow 0$$

and $$\bar{\epsilon} \longrightarrow \mathrm{RT} \text{ as } \mathrm{T}\longrightarrow\infty\,.$$

XC. *On the Refraction of Electro-Magnetic Waves in a Spherically Stratified Medium.* *By* Instructor-Captain T. Y. Baker, *R.N.* (*ret.*) *.

IN a material medium the velocity of electro-magnetic waves is affected by physical conditions such as density, degree of ionization, etc., that exist at the different points. It is generally held that in no case can the velocity of the wave-fronts exceed that of light *in vacuo*, so that the ratio of the velocity at any point in the medium to that *in vacuo* is a fraction. The inverse of this fraction is termed the refractive index at the point.

It is also held that energy is transmitted through a medium or through the æther in the form of transverse vibrations in which the wave-fronts move forward in directions always at right-angles to themselves, and the family of surfaces to which consecutive wave-fronts give rise are cut orthogonally by a set of curves which ordinarily are termed rays.

In considering the transmission of radiant energy it is in many cases more convenient to do so in terms of the rays rather than in terms of the wave-fronts.

As the different portions of the wave-fronts pass through different parts of a heterogeneous medium the velocities of those portions will change in accordance with the distribution of refractive index, giving rise to alterations in the shapes of the wave-fronts and consequently to the shapes of the rays.

A heterogeneous medium that is of special interest is one which is stratified, as far as refractive index is concerned, in concentric spherical layers. The earth's atmosphere approximates to something of this nature, though the change from night to day, altering, as it presumably does, the state of ionization, probably causes the stratification to be eccentric.

The consideration of the nature of the distribution of refractive index is not the only factor that governs the transmission of radiant energy through the atmosphere. Where the paths of the rays impinge upon the earth's surface there is possibly something in the nature of surface reflexion; where they pass from one layer into another where there is a rapid drop in refractive index there may be something analogous to what happens when a train of waves in the visible spectrum passes an air-glass surface. In such case part of the energy is transmitted from one medium to the

* Communicated by the Author.

other, part is reflected. For normal incidence the fraction of reflected energy is $\left(\frac{n-1}{n+1}\right)^2$ for oblique incidence the fraction is given by Fresnel's formula

$$\frac{1}{2}\left\{\frac{\sin^2(i-i')}{\sin^2(i+i')}+\frac{\tan^2(i-i')}{\tan^2(i+i')}\right\},$$

i and i' being the angles of the incidence of the ray with the normal in the two media.

There must also be considered, in order to bring the investigation into line with actual facts, what proportion of the energy is absorbed by the medium as the waves pass through it, and lastly, as the energy passes close to the surface of the earth, there must be considered the diffraction effects that are produced in the shadow.

All these considerations make the investigation of the paths of light-waves and wireless waves exceedingly complicated, and it is only the hypothetical cases in which some of the governing factors are omitted that rigorous mathematical treatment is possible.

In this paper such a course has been followed. Nothing but straight forward refraction is considered ; that is to say the waves are supposed to move everywhere at right-angles to themselves with velocities inversely proportional to the refractive index. All diffraction effects, all absorption effects, all surface reflexion effects, whether at the earth's surface, or possibly at the Heaviside layer, are left out of account, and the investigation which follows aims at showing what kind of physical effects in the reception of wireless signals, or in the distance and depression of the visible sea-horizon, will be found in this hypothetical atmosphere. No attempt is made to discuss the physical conditions such as density, degree of ionization, etc., at any point that produce the particular refractive index at that point, except that it is assumed that refractive index is a function of wave-length, and that consequently the distribution of refractive index with height is not necessarily the same for one wave-length as for another.

With those limitations there is investigated for this hypothetical atmosphere what must be the nature of the distribution of refractive index that will allow of a transmission of energy completely round the earth, and that will give rise to such phenomena as the skip distance. Within the region of the visible spectrum the refractive effects that give rise to a triple horizon, sometimes termed mirage, are considered.

The following nomenclature will be used throughout :—

r : the distance of any point from the centre of stratification.

p : perpendicular from the centre upon the tangent to the ray at any point.

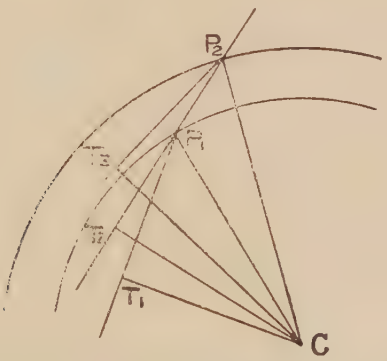
ϕ : the angle at any point between the tangent to the ray and the radius vector to the point.

n : the refractive index at any point.

R : the radius of the earth.

In fig. 1 are shown two surfaces separating media of refractive indices n_1, n_2, n_3 and a ray T_1P_2 is incident at the point P_1 making an angle T_1P_1C with the radius, *i. e.* with the normal at P_1.

Fig. 1.



The ray then passes through the layer of constant refractive index n_2 until it strikes the next interface at P_2, making an angle P_1P_2C with the normal, and is again refracted.

The points T_1, T_2, T_3 are the feet of the perpendiculars from the centre on the ray in each layer.

The equations of refraction are :

$$n_1 \sin T_1P_1C = n_2 \sin T_2P_1C,$$

$$n_2 \sin T_2P_2C = n_3 \sin T_3P_2C,$$

whence

$$n_1 T_1C/P_1C = n_2 T_2C/P_1C,$$

and

$$n_2 T_2C/P_2C = n_3 T_3C/P_2C,$$

or

$$n_1 T_1C = n_2T_2C = n_3 T_3C =, \text{ etc.}$$

And generally, if the layers are made infinitely thin, $n \,.\, T C$

is constant, the lines T_1P_1, T_2P_2, etc., becoming ultimately successive tangents to the curved path of the ray.

Consequently for a ray in a spherically stratified medium $np=\text{constant}=k$ say.

Hence, if n be known as a function of $r(=f(r))$, the equation of the ray is

$$np=pf(r)=k \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}$$

This is the general equation of any ray whatever in a spherically stratified medium.

The radius of curvature, ρ, being $\frac{rdr}{dp}$, it follows that

$$\rho=r/\frac{dp}{dr} \quad =-rn/p \ \frac{dn}{dr}.$$

In a particular case the waves in question may be waves of visible light travelling in a nearly horizontal direction over the surface of the sea. This will be the case when investigating the depression of the visible horizon from a point at a comparatively small height above sea-level. It is clear that in such case the rays must be very nearly horizontal for the whole of their length, and the ratio of r to p differs from unity only by very small amounts. The value of n for such light is known to be only very slightly greater than unity, and consequently the curvature of the ray at different points along its path is very nearly proportional to $\frac{dn}{dr}$. It is a customary practice in surveying to deal with the refraction by assuming that the ray is hollow towards the earth's surface and has a radius of about seven times the earth's radius. This is only consistent with what has been said above, provided the refractive index between sea-level and the point of observation diminishes at a uniform rate, a condition of affairs that is by no means always the case.

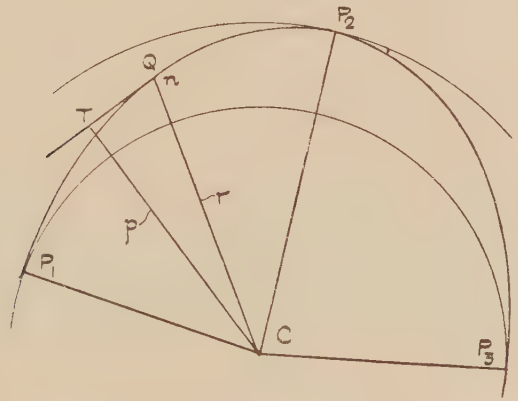
Reverting to equation (i.) suppose the ray is at some point P_1 (fig. 2) in its path travelling horizontally. Then at this point p_1 and r_1 are equal; and let the ray subsequently reach some higher level r_2 at a point P_2 where it is again horizontal. Then the equation $np=k$ gives rise to the relationship $n_1r_1=n_2r_2$, where n_1 and n_2 are refractive indices at P_1 and P_2.

The ray has travelled from P_1 to P_2 and clearly cannot have been horizontal for the whole of its length, since it has gone from a lower level to a higher one. At an intermediate point Q it must be travelling upwards, and since

$np = n_1 r_1 = n_2 r_2$ it follows that nr must be greater than $n_1 r_1$ or $n_2 r_2$ since $p < r$.

It follows, therefore, that in order that it may be possible for a ray to travel from P_1 to P_2 and be horizontal at both these places, the value of nr must increase to a maximum at some intermediate height between r_1 and r_2, and it is consequently of interest to examine the nature of the curve of nr plotted against r, but before doing so it should be noted that the subsequent track of the ray must from symmetry be along a track P_2P_3, which brings it back to the lower level r_1

Fig. 2.



where it is again horizontal. After P_3 it will ascend to the higher level r_2 and so on in succession. It is clear from this that such a ray never escapes out of the space between these two levels and that the requisite and sufficient conditions for this state of affairs are that nr should be greater in the intervening space than $n_1 r_1$ or $n_2 r_2$, which must equal one another.

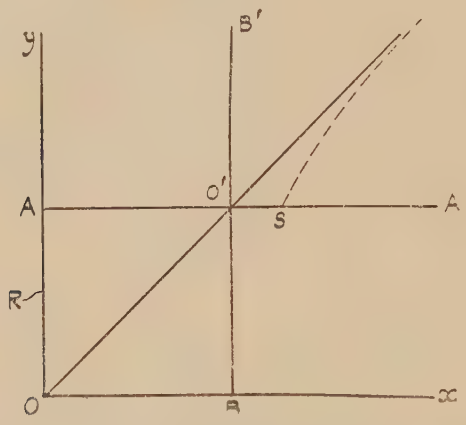
In fig. 3 is shown a curve of nr (abscissæ) plotted against r as ordinates.

In drawing such a figure it is to be remembered that the cases in contemplation are those of electro-magnetic waves in the earth's atmosphere at heights of 100 miles or so at the outside, and for refractive indices probably only slightly greater than unity.

It is quite certain that ultimately, when the height above the earth's surface is large enough, the value of n will be unity, and that consequently the curve for nr must be asymptotic to a line inclined at 45° with the axes.

In fig. 3 the pecked lines represent the value of nr, starting from the level AA′ where OA=R, the earth's radius.

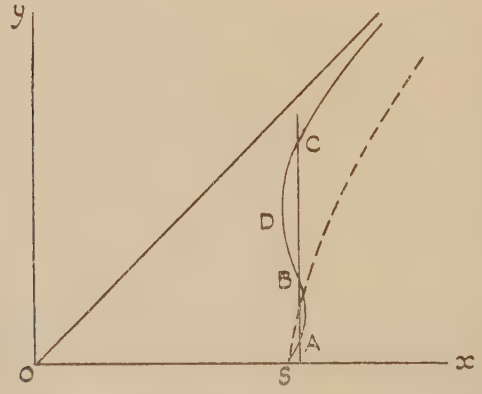
Fig. 3.



It is obvious that the curve can never cross the line OO′ at 45° through the origin, for if it did we should have a point at which the velocity was greater than the velocity of light *in vacuo*, which is contrary to the initial assumptions.

Hence the line OO′ becomes the asymptote to the curve.

Fig. 4.



The starting point S of the curve must be slightly to the right of O′ (for visible light $n=1{\cdot}0003$ approximately), and it is convenient at this point to re-draw the figure as fig. 4. taking O′A′ and O′B′ as axes and drawing everything on a much larger scale.

It is clear that in order to construct the track of any ray through the atmosphere, the value of n must be known in terms of the height above sea-level. This is known, of course, if the shape of the curve is known in fig. 4, and it is therefore of interest to consider the kind of curve that may be expected.

For visible light at all events $n-1$ drops uniformly with the density of the air, and as the density certainly diminishes as the height increases n becomes smaller and smaller and the curve gets continually closer to the asymptote.

The customary type of curve is that shown in the pecked line in fig. 4, but the curve shown in full line is not inconsistent with the assumption that n must diminish as the height increases. Whenever a curve of this shape occurs, then by drawing a line ABC parallel to OY to cut the curve in A, B, and C we find three separate levels for which nr has the same value. The figure shows that nr has the same value at A and B and is greater intermediately. The conditions are therefore suitable for a ray starting at r_1 (A) ultimately to reach a level r_2 (B), and thereafter to be confined entirely within the zone between these two heights.

On the other hand, a ray which is horizontal at a height r_2 cannot possibly reach the level r_3 of C, since intermediately nr has a lower value and np would be smaller still. There is therefore no possibility of a ray which is horizontal at the level r_2 getting across the zone r_2 to r_3, but if a ray starts off at a sufficient inclination it will do so.

In fig. 4, D is the point for which nr is a minimum $(=n_4r_4)$.

If a ray starts from the level r_1 at an inclination ϕ_1 with the vertical, then the constant of the equation of the ray is $n_1p_1=n_1r_1\sin\phi_1$.

If the ray just becomes horizontal at the level r_4 we have

$$n_1r_1\sin\phi_1=n_4r_4$$

or

$$\sin\phi_1=n_4r_4/n_1r_1.$$

Thus, any ray starting nearer to the horizontal than the above will become horizontal at a lower level than D and will return to lower levels. Any ray nearer the vertical will pass through the zone r_4, and with a curve for nr as shown in fig. 4 this ray will ultimately escape into the space beyond the earth's atmosphere.

The general conclusions to be drawn from the above statements are that if the curve for nr starts to the right from S, representing sea-level, no radiant energy can return

to sea-level unless the curve throws back to the left of the ordinate through S. Unless this throw-back exists radiation emitted at any inclination above the horizontal *must* escape from the atmosphere.

Now the general state of affairs with wave-lengths ordinarily used in wireless is that a considerable amount of such radiation does not escape but is retained within the atmosphere and is ultimately absorbed. With wireless waves, therefore, the general state of affairs is that the throw-back exists.

With the much shorter wave-lengths of visible light the conditions seem to be pretty well the reverse. The normal state of the nr-curve is that shown in fig. 4 in pecked line, but undoubtedly the throw-back does sometimes exist, for it seems impossible to explain by any other means the optical phenomenon of the triple horizon that is referred to later.

The shape of the curve shown in fig. 4 apparently furnishes an explanation of many of the phenomena of the reception of wireless signals.

It suffices to explain how a wave starting from a station can travel completely round the earth and be picked up again close to the transmitting station. It does not seem to be in the least degree necessary to suppose that such radiant energy travels round the earth with a series of alternate reflexions at the earth's surface and at the Heaviside layer. For suppose that the shape of the nr-curve in fig. 4 is a vertical straight line for the first three or four thousand feet. It is quite clear that a vertical plane-wave, say three thousand feet high, will, if it travels round the earth as a vertical plane-wave, have moved for the upper portion of the wave a distance $2\pi r$ and for the lower part $2\pi R$; and that since velocities at all points of the wave are inversely proportional to n the condition $nr =$ constant for the whole of the three thousand feet is a sufficient condition that wave-front will travel as a vertical plane-wave right round the earth.

Another phenomenon that can be explained by the shape of the nr-curve is that of there being frequently a blank space—the "skip distance"—over which there is a no reception of wireless waves. Signals are received at, say, 100 miles, and are then not picked up again until 1000 miles or more from the station.

In fig. 5 is shown a suitable nr-curve for such a result to be obtained. A is at the level of the transmitting station conveniently taken as sea-level. The curve throws back

and cuts the vertical through A at the point B (radial distance r_1).

We thus have $n_1 r_1 = n_0 R$ and nr intermediately greater than either of these values. After B, the curve subsequently becomes vertical again at E, re-crosses the vertical through A at C, and goes off asymptotically to the 45° line through the origin.

Fig. 5.

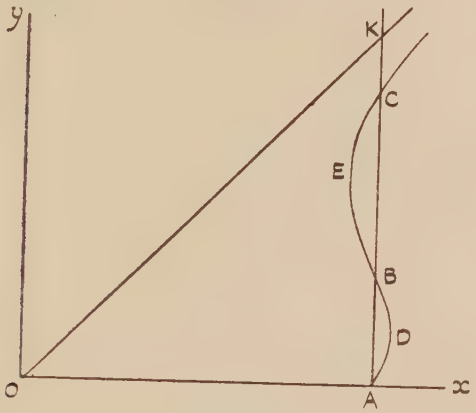
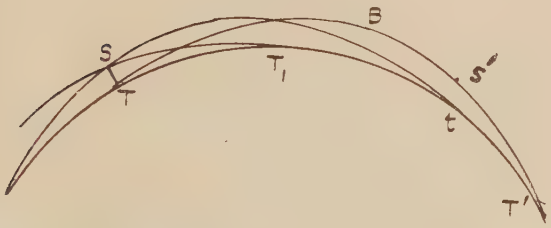


Fig. 6.



Consider now a ray starting from the station horizontally from T, fig. 6. That ray has an equation represented by $np = n_0 R$. Now

$$p = 1 \Big/ \left\{ u^2 + \left(\frac{du}{d\theta}\right)^2 \right\}^{\frac{1}{2}},$$

where $u = 1/r$. Hence

$$n = n_0 R \left\{ u^2 + \left(\frac{du}{d\theta}\right)^2 \right\}^{\frac{1}{2}},$$

and θ has the ordinary meaning when using polar co-ordinates, giving

$$\left(\frac{du}{d\theta}\right)^2 = \frac{n^2}{n_0^2 \mathrm{R}^2} - u^2,$$

or

$$\theta = \int \frac{n_0 \mathrm{R} du}{\{n^2 - u^2 n_0^2 \mathrm{R}^2\}^{\frac{1}{2}}}.$$

Consequently the ray reaches the sea-level again and is there horizontal at an angular distance θ, given by

$$\theta = z \int_{\mathrm{R}}^{r_1} \frac{n_0 \mathrm{R} dr}{r\{n^2 r^2 - n_0^2 \mathrm{R}^2\}^{\frac{1}{2}}},$$

an integrable equation if n is known in terms of the height.

Without an exact knowledge of the refractive index in terms of the height the integration cannot be carried out, but it is to be noticed that the denominator of the integrand, viz. $r\{n^2 r^2 - n_0^2 \mathrm{R}^2\}^{\frac{1}{2}}$ becomes smaller at any level as nr approaches $n_0 \mathrm{R}$.

Prescribing, therefore, that the ray is to reach the level of B, the greater the bulge between A and B in fig. 5 the closer to T will be the point T′ at which the ray returns to earth again. The flattening of the bulge extends the distance at which the wave returns to earth, but this process must not be pushed to the limit of making BA absolutely flat and vertical, for in such case the point where the throw-back of the nr-curve occurs is not at B but at a point immediately above A. In other words, with such an atmosphere a ray starting from A horizontally travels round the earth at sea-level, and becomes a particular ray of the plane vertical wave-motion previously considered. The ray-path TBT′ (fig. 6) is thus drawn for a definite throw-back height B (fig. 5) and a definite bulge ADB.

On the downward part of the wave-path there can be found a point S′ of a height above sea-level corresponding to the upper end S of the transmitting aerial.

Slide the ray-path backwards round the earth until S′ coincides with S and T′ then takes up a position T_1.

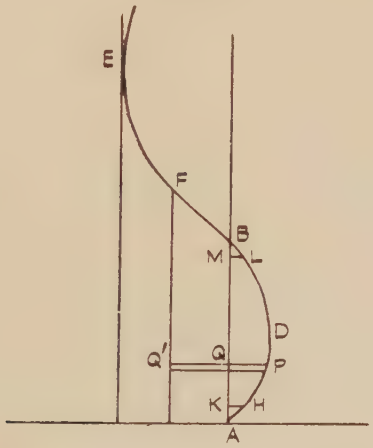
The atmosphere being stratified in concentric layers ST_1 is a possible ray-path, and a ray emitted from S in the suitable downward direction will reach sea-level at T_1. Any ray starting from S in a direction more nearly horizontal than ST_1 must have a larger value of p at the point S than the ray ST_1 has. Such a ray has consequently a larger constant np and cannot therefore reach sea-level at all.

The point T_1 is thus the farthest distance from the transmitting station at which the signals can be received. Beyond T_1 there will be no reception until some greater distance is reached. Slide the ray-path forward again until it passes through S and touches the sea-level the second time in t. This point t is a place at which a signal from S sent out in a suitable upward direction can be heard, and the space T_1t is a blank region (with a possible modification) over which there is no reception.

A ray starting from T in an upwardly-inclined direction has for the constant of its equation $n_0 R \sin \phi_0$, where ϕ_0 is the angle which the ray makes with the vertical at T. Let it be assumed that this inclination is only slightly above the horizontal, so that $\sin \phi_0$ is only very slightly less than unity.

Between B and E (fig. 7) it will be possible to find a point

Fig. 7.



F such that nr for this point, which is less than nr for B or nr for A, is equal to $n_0 R \sin \phi_0$. Thus the inclined ray starting from T in the direction ϕ_0 will become horizontal at the height of F and will afterwards return downwards and reach sea-level again at the inclination ϕ_0 with the vertical.

A similar integration formula can be applied to determine the place at which this ray returns, but the integration is extended over a bigger range of height.

The denominator of the integrand now becomes $r\{n^2r^2 - n_0^2R^2 \sin^2 \phi_0\}^{\frac{1}{2}}$, and for any particular value r the quantity under the square root has a greater value, according

to the figure, than the quantity $r\{n^2r^2-n_0^2R^2\}^{\frac{1}{2}}$ for the integration of the previous ray. Thus for the integration of this inclined ray the range of integration is greater but the integrand is less, and it cannot be generally said that the inclined ray comes down nearer to T than the horizontal ray, or *vice versa*.

As the inclination of the ray is increased the culminating point of the ray rises, but the limit is reached when F arrives at E where nr has a minimum value. Any inclination of the emitted ray nearer to the vertical gives, for the equation $np=K$, such a value of K that at no point in the atmosphere can the ray become horizontal, and the ray can therefore never return again to earth but will escape completely from the atmosphere.

It appears, therefore, from this reasoning

(i.) That transmission round the earth can only occur if the nr-curve throws back to smaller values than n_0R;

(ii.) that the critical emission direction is given by $\sin\phi=\frac{(nr)_{\text{min.}}}{n_0R}$. No ray emitted at a steeper inclination can ever return to earth, and the shape of the curve above E is absolutely immaterial so far as any of the received waves are concerned.

Explanations of wireless phenomena generally seem to be based upon some sort of idea that the ray received is either the "direct wave" as ST_1 at T_1 or the "reflected wave" as $S'T'$ at T', and usually in such explanation it is assumed that the reflected wave has gone up to the level of the Heavyside layer, generally taken to be from 70 km. to 120 km., and in some way reflected at that level. The possibilities of the waves reaching such heights must, of course, not be lost sight of, but it does not seem to be in the least degree necessary that they should all go so high. The points B and E might be so low and the points D and E so close to the line AB that reception would be possible over all parts of the globe without any ray travelling higher than a couple of kilometres—or even less.

It is not intended in this paper to do much more than offer a geometrical explanation of some phenomena of wireless transmission. It is recognized that such explanation cannot be complete; that the complications of diffraction, interference, and polarization must inevitably be present and

make the fitting of the parameters of the geometrical equations to observations somewhat uncertain. But in spite of these difficulties this geometrical aspect of wireless refraction seems to be reasonable, particularly when dealing with rays not too near the earth, and to lead to certain generalizations that are probably near the truth.

One of these generalizations is concerned with the type of value that n would necessarily have to have if the "reflected wave" had always to go to the height of the Heaviside layer, say to 300,000 feet above sea-level. At such a place p will equal 20,300,000 feet, taking 20,000,000 feet as a round number representing the earth's radius with sufficient accuracy. If the ray started off at an inclination ϕ_0 with the horizontal, we should have $n \times 20{,}300{,}000 = n_0 \sin \phi_0 \times 20{,}000{,}000$, so that $n_0 = 1{\cdot}015\ n \operatorname{cosec} \phi_0$.

Suppose now that it were possible for a ray to be emitted at an angle of 45° with the vertical and subsequently to return to earth, we should then have

$$n_0 = 1{\cdot}015 \times 1{\cdot}414\, n$$
$$= 1{\cdot}435\, n.$$

Now n, the refractive index at the culminating point, cannot be less than unity, so that the equation would involve a refractive index at the sea-level of at least 1·435, a figure probably far too great.

The fact is, unless one is prepared to accept quite high values for the refractive index at sea-level, the direction of emission of these rays which travel to the reception station must be confined to a few degrees above the horizontal, and it should be noted here that the higher the level of E in fig. 5 the greater must the sea-level value of n be.

Another generalization can be drawn from the integral which has already been discussed. The variable part is a denominator term frequently of very small dimensions. In order that this term can be computed with any degree of precision it is essential that the refractive index should be known with a precision that is greater still. To permit of that integral being evaluated with any semblance of reality the refractive index must be known to five or six decimal places and for a considerable number of levels up to the culminating level for the critical ray.

The values of n are not so known at the present time, and it seems almost impossible to obtain them by any direct means except possibly within the range of the visible spectrum.

Without such precision in the values of the refractive

index any attempt to compute the skip distance is farcical, but in order to illustrate the methods of doing so a fictitious atmosphere has been compiled which does not pretend to be any more than a blind guess at what the refractive index *might* be.

In making up this table of refractive indices regard had to be paid to ensuring a smoothness of values even when taken down to very small differences, and the easiest way of preserving this smoothness seemed to be to find a formula which when plotted had the suitable geometrical characteristics : that is to say, it had to represent a curve crossing the vertical at A, B, and C (fig. 5), and ultimately asymptotic to the line OK. Further, in making up the formula, regard had to be paid to the fact that it must be single valued in height, and that it should be as simple as possible to compute from.

The first consideration was the choice for n_0, which was taken to be 1·015. This figure allows the point E to be placed if required as high as 300,000 feet, but as will be seen later the point was made to come out considerably lower. In making up the formula it was convenient to make up the Cartesian equation of a curve in the form $x=f(y)$, for which the vertical and horizontal lines through A were taken as axes of coordinates. To reduce numbers to reasonable magnitudes the unit of length was taken to be one-thousandth part of the earth's radius, say 20,000 feet or about 6 kilometres. The radius of the earth being 1000, $n_0\mathrm{R}=1015$, and the asymptote OO′ of fig. 3 then cuts the axis $y=0$ at $x=-15$ and the axis $x=0$ at $y=15$.

Below this point, $y=15$, must come the three crossing points, and these were ultimately taken to be $y=0$, $y=4$, and $y=9$.

It was found that the equation

$$x=y-15+\frac{990}{y^2+2y+66}$$

was an equation which fulfilled the geometrical conditions, and from this formula the value of nr and n in the attached tables are computed.

Vertical tangents to this curve occur at $y=1{\cdot}60121$ where $x=0{\cdot}39601$, and at $y=6{\cdot}58740$ where $x=-{\cdot}33548$.

The value of nr along $x=0$ is 1015 and the minimum value of nr is 1014·66452.

Hence the critical angle of emission is given by

$$\sin\phi_0=1014{\cdot}66452/1015$$
$$=0{\cdot}99967,$$

making $\phi_0=88^\circ\,32'$.

With this atmosphere it is thus the fact that no rays emitted from sea-level at a greater inclination above the horizon than about $1\frac{1}{2}^\circ$ can be retained within the atmosphere, and the extreme height attained by any ray is round about 132,000 feet. Further, all rays emitted from the station which are retained within the atmosphere must culminate at heights lying between 80,000 feet ($y=4$) and 132,000 feet ($y=6{\cdot}58740$).

The equation of the ray emitted horizontally from sea-level is $np=1015$, and the distance TT′ (fig. 6) is given by

$$\theta = 2\int_{1000}^{1004} \frac{1015\,dr}{r\{n^2r^2-n_0^2\mathrm{R}^2\}^{\frac{1}{2}}}.$$

In evaluating this integral the factor r in the denominator varies between 1000 and 1004, so that a reasonably close approximation will be obtained if the integral is written

$$\int_{1000}^{1004} \frac{2{\cdot}026\,dr}{\{n^2r^2-(1015)^2\}^{\frac{1}{2}}}.$$

Again, the quantity $n^2r^2-(1015)^2$ has factors $nr-1015$ and $nr+1015$. The largest value of nr is 1015·39601 and the smallest 1015. The second factor can therefore be replaced without serious inaccuracy by 2030·15, whose square root is 45·057, and the integral becomes

$$\int_{1000}^{1004} \frac{\cdot 045057 dr}{(nr-1015)^{\frac{1}{2}}}.$$

A way of looking at this integral that conveys an idea of how it depends upon the shape of the curve can be derived from fig. 7. P is a point on the curve at distance r from the earth's centre, and PQ is $nr-1015$.

Hence the integral is $\cdot 045057\int \frac{dr}{\sqrt{\mathrm{PQ}}}$, the integration extending from A to B.

At both ends of the integration the integrand becomes infinite since the PQ becomes zero, and the greatest care must be taken when performing the integration by approximate methods.

One convenient method is as follows:—Cut off from the range of integration two parts, AHK and BLM, by lines KH and ML not too far away from A and B respectively, and treat the parts of the curve between A and H and B and L as straight lines.

Then the total integration can be broken into three parts:

(*a*) over the range A to K, (*b*) over the range K to M, and (*c*) over the range M to B.

For (*a*) we can transfer the variable to y measured vertically from A instead of r measured from the centre, and the integrand will be $\frac{1}{\sqrt{my}}$, where $m = \tan \mathrm{HAK}$. We thus have over the area AHK:

$$\text{Integral} = \cdot 045057 \int_0^{\mathrm{AK}} \frac{dy}{\sqrt{my}},$$

$$= \frac{\cdot 045057}{\sqrt{m}} \int_0^{\mathrm{AK}} \frac{dy}{\sqrt{y}},$$

$$= \frac{\cdot 045057}{\sqrt{m}} \times 2\sqrt{\mathrm{AK}},$$

$$= \frac{\cdot 090114}{\sqrt{m}} \sqrt{\mathrm{AK}}.$$

In a similar manner, the part (*c*) of the integrand

$$= \frac{\cdot 090114}{\sqrt{m'}} \sqrt{\mathrm{BM}},$$

where m' is the corresponding slope of the tangent at B.

It will be seen from this that as the inclination to the vertical of the tangents at A and B becomes smaller the integrals (*a*) and (*c*) become larger, and that if the curve either at A or at B touched the vertical AB the integral would become infinite. For the central portion between K and M the integrand nowhere becomes infinite, and ordinary methods of approximate integration are suitable.

The values of nr for $=1000$, 1001, 1002, 1003, 1004 are 1015, 1015·034782, 1015·38018, 1015·22222, and 1015; and it will be seen from these figures that one must have accuracy in the first three decimal places to obtain accuracy in the integration. When the value of r has been divided into nr it becomes immediately clear that the determination of the point T′, at which the ray TBT′ of fig. 6 returns to sea-level, involves an accurate knowledge of the fifth and sixth decimal places in the value of n.

It is important that this point should be made quite clear. *It is quite impossible to attempt to compute the track of a ray through the atmosphere for an electro-magnetic wave that starts off in a known direction from a known point unless the refractive index for all levels that it will reach is known with accuracy to about the fifth decimal place.* Whether it is possible to

measure the refractive index with that accuracy does not form part of this paper.

The values of $nr-1015$ near the point of A are:

r.	$nr-1015$.
1000	0
1000·1	0·0524
1000·2	0·1007
1000·3	0·1448
1000·4	0·1849
1000·5	0·2212

At the other end the values of $nr-1015$ are:

r.	$nr-1015$.
1003·5	0·1129
1003·6	0·0902
1003·7	0·0675
1003·8	0·0449
1003·9	0·0223
1004·0	0·0000

The figures show that the curves are reasonably straight at the ends for a space of 0·2, and that the values of m and m' are approximately 0·55 and 0·22.

Integrating the middle portion by Simpson's Rule, w obtain:

$$\begin{aligned} \theta(a) &= 0\,0544 \\ \theta(b) &= 0{\cdot}3353 \\ \theta(c) &= 0{\cdot}0859 \\ \text{Total} &= 0{\cdot}4756 \text{ or } 1635 \text{ sea miles.} \end{aligned}$$

Hence a ray starting horizontally at sea-level reaches sea-level again at a point 1635 miles away.

Take, for purposes of illu-tration, a height ST (fig. 6) of 500 feet, which is in the units we are using 0·025.

The angular distance TT, as obtained by integration of the same formula between l mits 1000 and 1000·025. Integration here must be of the type (a), and the value is $\dfrac{\cdot 090114\times\sqrt{\cdot 025}}{\sqrt{0{\cdot}55}}=\cdot 0192$, or 65 sea miles.

Consequently the distance T_1t of fig 6 is 1505 miles, and this will be the skip distance unless any of the inclined rays starting from the transmitting station descend within this space.

All the rays emitted from the station culminate within the space $r=1004$ to $r=1006{\cdot}5874$.

The point of descent for an inclined ray is given by

$$\theta = 2\int_{R}^{r} \frac{Rn_0 \sin\phi_0 dr}{r\sqrt{n^2r^2 - n_0^2R^2\sin^2\phi_0}},$$

where the range of integration extends at the most from 1000 to 1006·5874.

The same preliminary approximation, involving treating r and $\sqrt{nr + n_0R\sin\phi_0}$ as constants, can be made and the equation then becomes

$$\theta = \int_{1000} \frac{\cdot 045057 dr}{(nr - 1015\sin\phi_0)^{\frac{1}{2}}}.$$

The angle ϕ_0 is in any case greater than $88\frac{1}{2}°$. No appreciable loss of accuracy is entailed by writing $\sin\phi_0$ as unity, except in the denominator term.

Written in this way it is clear that the integral can be interpreted by means of fig. 7.

A ray which culminates at F will have a value for ϕ_0 easily calculated, and $nr - n_0R\sin\phi_0$ now becomes the distance PQ′.

The altered limits of integration obviously can affect the integral, either to make it larger or to make it smaller.

The increase in the length from PQ to PQ′ is a reducing factor but the increase in range of integration tends to make the value larger, and especially should it be noted that the shape of the nr-curve shows that the value m' in (*c*) is becoming smaller as the culminating height becomes greater. Ultimately the reduction in the integral due to increased values of PQ′ must be overbalanced by the increase due to the end values, and finally when F reaches E the integral will be infinite.

The maximum height (E) which any rays starting from the earth's surface can reach and yet return to the earth's surface is a circular asymptote, and the ray must in theory travel an infinite distance before it starts to descend and another infinite distance before it reaches sea-level again. In the neighbourhood of E nr is sensibly constant, which is the condition for a vertical wave travelling at constant height round the earth.

From the above reasoning it can be seen that as the inclination above the horizontal of a ray starting from T is increased, the point of descent to the earth's surface may move from T′ towards T or may move away from T. It is entirely a question of the distribution of refractive index with height, and it is extraordinarily sensitive to small

changes in the refractive index distribution. If the point moves towards T it may even go as far as T_1 and wash out the whole of the skip distance; or it may go part of the way and diminish the skip distance T_1t that was found for rays horizontal at sea-level. The skip distance is therefore limited by T_1 at the one end, and at the other by either t or some point at which an inclined ray comes down.

But in any case, as the inclination gets larger and larger and approaches the critical angle, the *point of descent must ultimately pass beyond* T′ *and cover continuously the whole of the earth's circumference beyond that point.*

There is thus no second skip distance, as there would be if reception were confined to those rays that are horizontal at the earth's surface.

Let us now investigate whether the faked atmosphere has a skip distance or not.

It will be convenient to integrate

$$\int_{1000} \frac{\cdot 045057 dr}{(nr - 1015 \sin \phi_0)^{\frac{1}{2}}}$$

up to limits 1004·5, 1005, 1005·5, 1006, and 1006·5.

In theory the minimum distance of the point of descent could be obtained by the calculus of variation, but this would involve the computation of $\int \frac{dr}{PQ^{\frac{3}{2}}}$, and as PQ is supposed to be known only in terms of numerical values the labour involved would be too great.

It is therefore preferable to plot a curve for the five ranges given above and determine the minimum by inspection of the curve.

The values of x for

$$x = y - 15 + \frac{990}{y^2 + 2y + 66},$$

when y has values 4·5, 5·0, 5·5, 6·0, 6·5, are

$$-0{\cdot}106299, \quad -0{\cdot}198020, \quad -0{\cdot}269231, \quad -0{\cdot}315784, \quad -{\cdot}335052,$$

and the values of nr are

$$1014{\cdot}893701, \quad 1014{\cdot}801980, \quad 1014{\cdot}730769, \quad 1014{\cdot}684216, \quad 1014{\cdot}664948.$$

Integrating the equation

$$\theta = \int_{1000} \frac{\cdot 045057 dr}{\{nr - n_0 \mathrm{R} \sin \phi_0\}^{\frac{1}{2}}}$$

to the appropriate upper limit, and with the appropriate values of $n_0 R \sin \phi_0$, we obtain the following values :—

Inclination of ray.	Range to point of descent.	Height of culminating point.
$\phi = 90°$	= 1635 miles.	80,000 feet.
89° 10′	1585 „	90,000 „
88° 53′	1693 „	100,000 „
88° 41′	1920 „	110,000 „
88° 34′	2270 „	120,000 „
88° 33′	3370 „	130,000 „
88° 32′	——	132,000 „ approx.

It will be seen that the horizontal ray is not the one which determines the skip distance, the point of descent initially approaching the transmitting station as the inclination is raised above the horizontal direction, and the turning point is reached at about 1580 miles. The inclination of this ray is about 0° 45′ above the horizontal, and the culminating point at about 87,000 feet. Thereafter the point of descent recedes, and consequently all points beyond 1580 miles from T can receive rays starting from T. Rays starting from S will be received about 70 miles nearer.

Hence the near reception will extend to about 65 miles, and the skip distance will be about 1450 miles.

Examination of the figures in the above integration leads to the conclusion that only a very small alteration in the shape of the nr-curve is required to wash out the whole of the skip distance. Give the part ADB (fig. 5) a little more bulge, and the point of descent for the horizontal ray is brought closer in. The increased bulge naturally leads to the curve crossing the vertical at B more squarely, and this in its turn will reduce the end values of the integrals for the inclined rays. Thus it will be seen that only quite small alterations in the refractive index distribution are necessary to cut out the skip distance altogether and give continuous reception all the way from the starting point.

Now presumably the refractive index distribution depends upon the wave-length, much in the same way as a piece of glass has different refractive indices for the red and blue ends of the spectrum. There is then no impossibility in the atmosphere giving rise to a skip distance for one wave-length and not for another.

One further point should be noted. If the nr-curve started

off to the left of the vertical instead of the right, as it would do if the point B (fig. 5) was at the level of the earth's surface, then no ray starting horizontally would get any distance at all; there would be no skip distance but continuous reception by inclined rays right away from the start.

There would still be a critical angle of emission given by $\sin\phi_0 = \frac{(nr)_{\min.}}{n_0 R}$, and only rays starting off below this direction could possibly return to earth again.

In the fictitious atmosphere that has been used to illustrate the methods of integration the points B and C were taken at the arbitrary heights of 80,000 and 180,000 feet and, resulting from the formula by means of which the table of values for n was compiled, all the rays by which reception takes place culminate between heights of 80,000 and 132,000 feet. There is no reason to assume that these two heights are in any way connected with the Heaviside layer. It would, in fact, have been just as easy to make up an atmosphere in which B and E both lay below 10,000 feet, and in which wireless transmission would take place along rays which did not approach anywhere near the Heaviside layer. It should be noted also that the higher the point C is placed the greater the need for having a refractive index at the earth's surface which is considerably more than unity. Conversely the smaller the refractive index at the earth's surface the lower must be point C if the wireless wave is to be bound to the atmosphere.

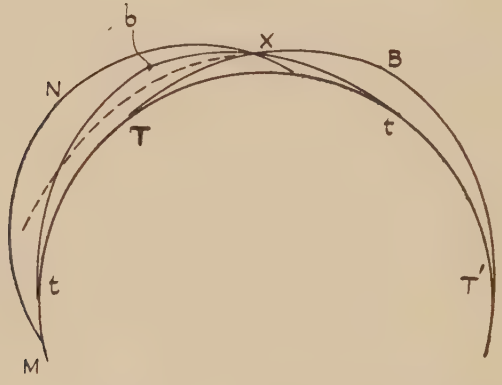
No attempt is made in the above work to explain the physical causes of such a distribution of refractive index, nor has any account been taken of the reflexion and diffraction of the waves by the earth's surface or of the scattering of energy by the medium. Such physical effects certainly exist. They complicate the problem of determining the transference of radiant energy, but it does not seem to be imperative to take account of them in order to have a reasonable explanation of most of the phenomena of wireless transmission and reception.

The shape of the curve of nr has important bearings also on the passage of visible light through the atmosphere. The value for n at the earth's surface is much lower (1·000293) and the phenomenon of the earth-bound ray has to take place, if at all, within a hundred feet or so. In point of fact, the normal case for the refractive index of visible light is that of fig. 3, and it is only in exceptional cases that a throw-back of the curve occurs to the left of the vertical. When such is the case, a ray of light which is horizontal at

T (fig. 8) becomes horizontal again at B, and comes down again to the earth's surface at T′, the distance TT′ depending as before upon the value of the integral $\int_R^r \frac{dr}{\{nr - n_0 R\}^{\frac{1}{2}}}$ up to the level of B.

Now the height AK up to the asymptote (fig. 5) is only 0·293 in the units used, *i.e.*, something under 6000 feet, but the point B must be much lower than that. Probably it is seldom more than a couple of hundred feet above sea-level. A man viewing the sea-horizon from the point B

Fig. 8.



will receive the rays truly horizontally and the earth will to him appear flat.

For observations made between T and B the horizon will appear depressed, the maximum amount occurring at D where D is at the height of the point D in fig. 5.

Now, through any point X on TB an exactly similar curve *tbt′* could be drawn which started horizontally at *t* became horizontal again at *b* and at X was moving downwards. Hence the man at X will receive two rays, one coming to him inclined upwards and another inclined equally downwards, both of which have started tangentially to the earth's surface.

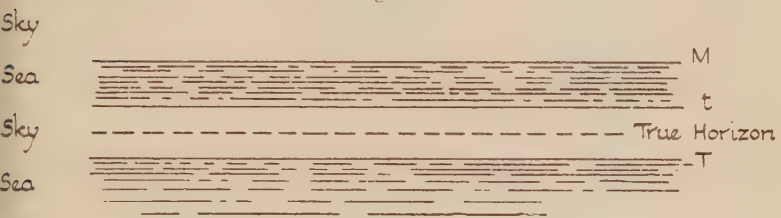
For each of them the equation is $pn = n_0 R$, and since n is the same for both rays at the point X (radial distance r) the values of p must be the same at this point, this giving equal inclinations above and below the horizontal.

Any ray which reaches X between these two directions

must by the geometry of the figure have a larger value for p, and consequently a larger value for np. Thus the constant in the equation $np = K$ for this ray is greater than $n_0 R$, and by the assumption as regards the shape of the nr-curve for the atmosphere this value is reached for some height that is above the earth's surface.

The observer will thus see, as his eye travels upwards from the sea below him, sea up to the direction XT. There will there be a sharp dividing-line separating sea from sky and a space of clear sky between the direction XT and the direction Xbt. Above that direction he sees sea again, and this continues until the direction is reached in which a ray having started off at the critical angle has made an infinite number of revolutions round the earth to the circular asymptote, and subsequently another infinite number until it reaches his eye on the downward path. Actually, of course,

Fig. 9.



it does no such thing, the absorption by the atmosphere being far too great, so that the limit is set not by the critical angle but by the range of visibility, and the third line which separates sky from air is an indistinct one corresponding to a ray inclined at M, the extreme range of visibility which culminates at N higher than b or B and reaches his eye in a downward direction.

The appearance of the sea-horizon as seen from X is thus that shown in fig. 9, and the true horizontal occurs half-way between the lines t and T and is shown in pecked line.

It is hardly necessary to compile a table of refractive indices that will serve to compute the angular distance between T and t. It is sufficient to state that owing to the very low refractive index for visible light the heights of B and E in fig. 5 can only be of the order of a hundred feet or so, and that in consequence the distances from X of T, t, and M in fig. 8 can be only a few miles, and the angular separation of the three horizons only a matter of a minute or two at the most. The phenomenon of the triple horizon

is well known to anyone who has had to take astronomical observations for navigation purposes. It is generally referred to as mirage without any very satisfactory explanation of how it is caused. Not infrequently the navigator assumes that the detached strip of sea is a layer of low-lying haze or cloud.

As the point X moves upwards towards B the two horizons *t* and T merge into one another and at the point B have just disappeared. It is therefore wrong to say, as was done above, that the observer at B sees the earth as flat. He sees it hollow, the line dividing sea from sky being that of M, which is the indistinct line coming from the extreme range of visibility. As the observer goes higher still he ultimately reaches a height at which he is at the culminating point of an inclined ray which started from the extreme range of visibility. He then has a "dip" value of zero and for greater heights the sea-horizon is depressed below the horizontal. It appears, therefore, that a formula for the dip proportional to the square-root of the height as is taken in all mathematical tables is quite wide of the mark when the particular atmospheric conditions occur that give rise to a throw-back in the *nr*-curve.

It is of course known that cases occur in which the dip tables are in error by considerable amounts, although on the average they probably are near enough. In a series of observations carried out some years ago by an American naval officer off the coast of California, errors of the dip tables up to nine minutes were observed. In actual practice a few minutes error in measurement of the sun's altitude results in an equal number of miles error in determination of the ship's position; and as it is only when out of sight of land that astronomical observation becomes necessary the safety of a ship is never endangered. For surveying work, or in cable-laying ships, however, the matter is different, and the only way to avoid the dip errors is either to observe several stars in different azimuths or to measure the dip at the time by special apparatus which measures the angular distance from the front to the back horizon.

Table of Values for Refractive Index.

In this table the unit of length is taken to be 1/1000 part of the earth's radius, approximately 20,000 feet or 6 kilometres. The formula used to construct the "*nr*"-curve is $x = y - 15 + \frac{990}{y^2 + 29 + 66}$, being the height above the earth's surface, and $nr = 1015 + x$.

TABLE.

y.	nr.	n.
0	1015·0	1·015
0·1	1015·05242	1·0149509
0·2	·10066	1·0148977
0·3	·14480	1·0148403
0·4	·18495	1·0147790
0·5	·22119	1·0147138
0·6	·25364	1·0146448
0·7	·28241	1·0145722
0·8	·30762	1·0144960
0·9	·32938	1·0144164
1·0	·34783	1·0143335
1·1	·36307	1·0142474
1·2	·37526	1·0141583
1·3	·38451	1·0140662
1·4	·39096	1·0139714
1·5	·39484	1·0138740
1·6	·39599	1·0137740
1·7	·39484	1·0136716
1·8	·39143	1·0135670
1·9	·38590	1·0134603
2·0	·37839	1·0133517
2·1	·36900	1·0132412
2·2	·35789	1·0131290
2·3	·34520	1·0130153
2·4	·33103	1·0129000
2·5	·31553	1·0127835
2·6	·29882	1·0126659
2·7	·28101	1·0125471
2·8	·26224	1·0124274
2·9	·24260	1·0123069
3·0	·22222	1·0121856
3·1	·20121	1·0120638
3·2	·17967	1·0119415
3·3	·15771	1·0118187
3·4	·13542	1·0116956
3·5	·11290	1·0115724
3·6	·09025	1·0114491
3·7	·06755	1·0113257
3·8	·04489	1·0112023
3·9	·02235	1·0110792
4·0	·00000	1·0109562
4·1	1014·97793	1·0108335
4·2	·95619	1·0107112
4·3	·93482	1·0105893
4·4	·91402	1·0104679
4·5	·89370	1·0103471

TABLE (*continued*).

y.	*w*.	*n*.
4·6	1014·87397	1·0102270
4·7	·85489	1·0101074
4·8	·83650	1·0099886
4·9	·81885	1·0098705
5·0	·80198	1·0097532
5·1	·78594	1·0096368
5·2	·77077	1·0095213
5·3	·75649	1·0094066
5·4	·74315	1·0092930
5·5	·73077	1·0091803
5·6	·71938	1·0090686
5·7	·70901	1·0089580
5·8	·69968	1·0088484
5·9	·69140	1·0087398
6·0	·68422	1·0086324
6·1	·67811	1·0085261
6·2	·67530	1·0084231
6·3	·66924	1·0083169
6·4	·66653	1·0082139
6·5	·66495	1·0081123
6·6	·66452	1·0080116
6·7	·66524	1·0079122
6·8	·66713	1·0078140
6·9	·67019	1·0077170
7·0	·67442	1·0076211

NOTE.—The figures given above are *not* measured values of the refractive index but are computed from a formula. The only justification for carrying the figures to a number of decimal places much in excess of anything that has ever been measured lies in the explanation given on p. 970.

The thanks of the author are due to the Admiralty for permission to publish this paper.

XCI. *The Adsorption of Butyric Acid on Water Surfaces.* *By* C. R. BURY *.

SEVERAL authors have already calculated from surface-tension measurements the amount of butyric acid adsorbed on water surfaces: in the absence of activity data, they have assumed the activity of butyric acid to be proportional to its concentration, or to the concentration of its unionized molecules. Recent measurements of the freezing-points of butyric-acid solutions (Jones and Bury, Phil. Mag. (7) iv. p. 843, 1927) have shown this assumption to be unjustifiable; and it therefore seems desirable to

* Communicated by the Author.

recalculate the adsorption, making use of the activities deduced by these authors.

The most interesting surface-tension measurements are those of the air-solution surface by Drucker (*Zeit. Phys. Chem.* lii. p. 649, 1905): though possibly less accurate than more recent measurements by Szyszkowski (*Zeit. Phys. Chem.* lxiv. p. 395, 1908), they cover the whole range from pure water to pure acid. In Table I. are given the concentrations of butyric acid in gram molecules per 1000 grams solution (C) and their relative surface tensions at 25° (σ/σ_0), taken from his paper. From these have been calculated the concentrations in gram molecules per 1000 grams water (m) and the absolute surface tensions in dynes per cm. (σ), using Volkmann's value for pure water (Wied. *Ann.* lvi. p. 457, 1895). Activity coefficients (γ) have been obtained by interpolation from the data of Jones and Bury, and from these the activities ($a=\gamma m$).

Table I.

C.	log C.	γ.	m.	a.	log a.	σ/σ_0.	σ.
0			0			1·000	71·78
0·01583	−1·8005		0·01585			0·963	69·12
0·03561	−1·4485		0·03571			0·900	64·60
0·08247	−1·0837		0·08308			0·828	59·43
0·1187	−0·9255		0·1199			0·778	55·85
0·2675	−0·5727	1·149	0·2740	0·3148	−0·5020	0·659	47·30
0·4353	−0·3612	1·117	0·4527	0·5057	−0·2961	0·577	41·42
0·9802	−0·0087	0·983	1·073	1·055	+0·0231	0·456	32·74
2·834	+0·4524	0·454	3·777	1·715	+0·2342	0·389	27·92
6·920	+0·8401	0·100	17·72	1·772	+0·2485	—	(27·0)*
9·015	+0·9549		43·76			0·374	26·85
11·38	+1·0561					0·361	25·91

* Interpolated from the γ-log C graph.

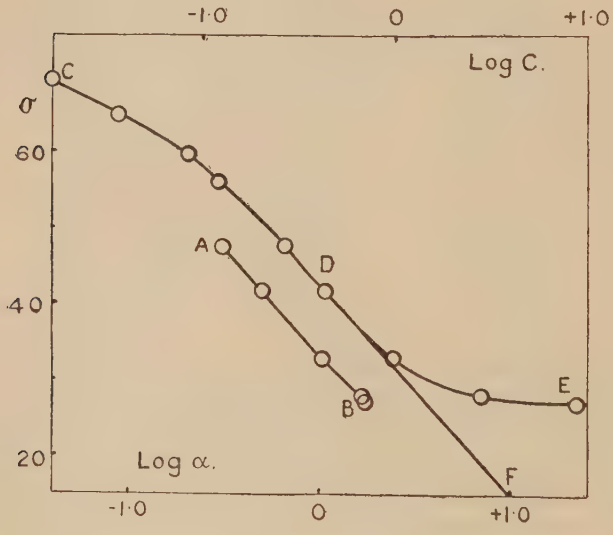
The line AB in the figure represents Drucker's surface tensions plotted against log a: within the limits of experimental error, it is straight over the whole range of concentrations (0·27–17·7 m.) for which activity data are available. From Gibbs's equation, which can be written in the form

$$\Gamma = -\frac{1}{2{\cdot}303\mathrm{RT}}\frac{d\sigma}{d\log a},$$

where Γ is the amount of butyric acid adsorbed at the

water-air surface as defined by Gibbs (Scientific Papers, 1906, vol. i. p. 234), R is the gas constant, and T the absolute temperature, it follows that Γ is proportional to the slope of this line, and is therefore constant over the whole range. Its numerical value is $4{\cdot}75 \times 10^{-10}$ gram molecules per $(\text{cm.})^2$, which corresponds to an area of $34{\cdot}7 \times 10^{-16}$ $(\text{cm.})^2$ per molecule. There is a possible error of about 2 per cent. in these figures, due to the graphical method of calculation employed.

Fig. 1.



Substances which form monomolecular layers or are strongly adsorbed at the surface of a solvent appear always to form very abnormal solutions in that solvent; it is therefore dangerous to use any but experimentally determined activities in calculations of this type. Thus, if activity was assumed to be proportional to concentration, Γ would be proportional to the slope of the line CDE in the figure, which represents surface tensions plotted against log C. Adsorption would then increase with concentration up to the point of inflexion (D) at 0·35 per gram molecules per 1000 grams solution, and then decrease. There would possibly be a small range of concentrations in the

immediate neighbourhood of the point of inflexion over which Γ was constant, but this could hardly be regarded as evidence of the formation of a saturated surface layer.

Szyszkowski (*loc. cit.*) found that the relation between surface tension and concentration of solutions of certain fatty acids could be expressed by an equation of the type

$$\sigma = \sigma_0 - 0{\cdot}411\sigma_0 \log\left(\frac{C}{0{\cdot}0475}+1\right),$$

where σ_0 is the surface tension of pure water. The numerical constants have been evaluated by Szyszkowski from Drucker's experimental data for butyric acid at 25°. This equation is represented by the line CDF in the figure; it obviously fails to represent the experimental data at high concentrations. Langmuir's well-known argument as to the formation of monomolecular surface layers by soluble fatty acids (J. Amer. Chem. Soc. xxxix. p. 1883, 1917) is based on two assumptions: that activity is proportional to concentration, and that Szyszkowski's equation is valid at high concentrations. The errors introduced by the two assumptions practically cancel out: *i. e.*, the lines AB and DF on the figure are almost parallel.

A number of calculations by previous investigators and by myself are summarized in Table II., where Γ is the maximum adsorption and A the minimum area occupied by a molecule of butyric acid. My own calculations have been made by the method described above, using the activities of Jones and Bury. The activities are strictly valid only at the freezing-point; their use at higher temperatures must lead to underestimation of A, but this is probably negligible. It is therefore probable that the increase of A with temperature, which is obvious from the measurements of Drucker and Szyszkowski, is real, and not entirely due to the errors of experiment and calculation. The difference between the results of these two observers is attributed by the latter to the greater purity of his acid.

A saturated surface layer does not appear to us to be formed at a benzene-water interface. The curve obtained, even when activities are used, has the same form as the curve CDE in the figure. In the absence of any information in the original paper as to the temperature at which the surface-tension measurements were made, I have assumed, in my calculation, that this was 20° C.

TABLE II.

Surface.	Surface-tension measurements by	T.	$\Gamma \times 10^{10}$.	$A \times 10^{16}$.	Calculated by
Water-air	(Szyszkowski's equation.)			31	Langmuir *.
,, ...	?	?	4·53	36·3	Harkins & King †.
,, ...	?	?		32	Harkins & McLaughlin‡.
,, ...	Drucker *.	25°	4·75	34·7	Bury.
,, ...	,,	35°	4·68	35·3	,,
,, ...	Szyszkowski *.	2–3°	5·45	30·3	,,
,, ...	,,	18–19°	5·34	30·9	,,
,, ...	,,	25°	5·16	32·0	,,
Water-benzene .	Harkins & King †.	?	4·60	35·9	Harkins & King †.
,,	,,			32	Harkins & McLaughlin‡.
,,	,,		4·89	33·8	Bury.
Water-hexane...	Harkins & McLaughlin ‡.	20°		32	Harkins & McLaughlin‡.
, ...	,,		5·97	27·6	Bury.

* *Loc. cit.* † J. Amer. Chem. Soc. xli. p. 970 (1919).
‡ J. Amer. Chem. Soc. xlvii. p. 1610 (1925).

In conclusion, butyric acid forms saturated surface layers at the water-air surface, in which the area occupied by a molecule of acid is independent of concentration over wide ranges, but appears to increase slightly with temperature. The area occupied by a molecule at the water-hexane interface is appreciably smaller than at the water-air surface, but still larger than that occupied by the higher, insoluble, fatty acids and alcohols. There is no evidence of the formation of a saturated surface layer at the water-benzene interface.

Edward Davies Chemical Laboratories,
University College of Wales,
Aberystwyth.
August 22nd, 1927.

XCII. *The Structure of the Isomorphic Substances* $N(CH_3)_4I$, $N(CH_3)_4Br$, $N(CH_3)_4Cl$. *By* L. Vegard, *Dr. Philos., Professor of Physics, Oslo, and* Karl Sollesnes, *Research Student at the Physical Laboratory, Oslo* *.

[Plates XX. & XXI.]

Introduction.

§ 1. The structure of the tetramethylammonium iodide ($N(CH_3)_4I$) was treated by one of us in two previous papers.

In a paper published in 1917 † the structure was investigated by the Bragg ionization method. The dimensions of the unit cell were determined and found to contain two molecules, and the atoms were found to be arranged in a way which corresponds to the space group D^7_{4h}.

In a paper recently published ‡ by one of us, together with T. Berge, results were given of investigations on the structure of $N(CH_3)_4I$ by means of the powder method. The main results stated above as to the magnitude of the unit cell and the space group of the atomic arrangement were confirmed, but in order to obtain a satisfactory agreement between observed and calculated intensities of the powder diagrams, certain changes of the parameters had to be made.

In $N(CH_3)_4I$ the reflecting power of nitrogen is very small as compared with that of iodine, and consequently the intensities of the spectral lines are very little influenced by the position of nitrogen atoms

From the investigations of the iodine compound alone, we are left in some uncertainty as to the position of the nitrogen atoms. The influence of these atoms, however, should be more prominent in the isomorphic compounds of bromine and chlorine, and for this reason it was of great importance to extend the investigations also to these two substances.

It appeared to be somewhat difficult to obtain good powder diagrams from the chlorine compounds, because it was very hygroscopic. The difficulty was overcome by introducing the powder into a capillary tube covered with a glass cap, which could be removed after the tubes were put into the X-ray camera. In this way we finally obtained good spectrograms for both the Br and Cl compound. Reproductions of the films are given in the plate (Pl. XX.).

* Communicated by the Authors.

† L. Vegard, "Results of Crystal Analysis.—IV." Phil. Mag. xxxiii. p. 395 (1917) (I.).

‡ L. Vegard & T. Berge, *Det. Norske Vid. Akad. Skr.* i. No. 10 (1926) (II.).

Specification of the Lattice.

§ 2. The lines appearing on the film were easily identified from the analogy with the iodine compound already analysed. For all three compounds the unit cell contains two molecules. (See Tables II. and III.)

In order to find accurate values for the dimensions of the lattice, the powder was mixed with some rock-salt, and from the rock-salt lines the correction terms for the film were found. A number of the stronger lines were carefully measured, and the values of a and c calculated by the method of least squares. The results are given in Table I.

TABLE I.

	a.	c.	c/a.
$N(CH_3)_4I$	7·91 Å.	5·74 Å.	0·7256
$N(CH_3)_4Br$	7·708 „	5·501 „	0·7137
$N(CH_3)_4Cl$	7·588 „	5·374 „	0·7082

The intensities of the observed lines as they were estimated from the spectrograms are given in Tables II., III., and IV., and in figs. 1 and 2 (A) for the Cl and Br compounds respectively. For the sake of completeness, a similar diagram is given in fig. 3 for the iodine compound.

TABLE II.—$N(CH_3)_4Cl$.

No.	$h_1 h_2 h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)^2$.	I obs.	I cal.
	1 0 0	1			
1	0 0 1	1·99			
	1 1 0	2·00			
2	(1 0 1) β				
3	1 0 1	2·99	3·026	8	9·6
4	(1 1 1) β	(3·31)	(3·33)		
5	1 1 1	3·99	4·009	100	100
	2 0 0	4·00			
	2 1 0	5·00	—	—	0
6	(2 1 1) β	(5·79)	5·79		0·1
7	2 0 1	5·99			
8	2 1 1	6·99	7·00	70	53·2
9	(1 0 2) β	(7·43)	7·37		
10	0 0 2	7·97	7·938	4	4·7
	2 2 0	8·00			
11	(2 2 1) β	(8·27)	(8·28)		

TABLE II. (*continued*).

No.	$h_1\ h_2\ h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)^2$.	I obs.	I cal.
12	1 0 2	8·97	8·951	28	20·1
	3 0 0	9·00	—	—	0
13	2 2 1	9·99	9·993	20	16·4
	3 1 0	10·00			2·0
	3 0 1	10·99			
14	2 0 2	11·97	11·991	20	13·7
	3 1 1	11·99			
15	2 1 2	12·97	12·964	5	2·5
	3 2 0	13·00	—	—	0
16	(3 0 2) β	(14·05)	14·01		
17	3 2 1	14·99	14·936	5	3
18	2 2 2	15·97	15·91	8	10·3
	4 0 0	16·00			
19	3 0 2	16·97	16·974	28	32
	4 1 0	17·00	—	—	0
20	0 0 3	17·94	17·993	16	17
	3 1 2	17·97			
	4 0 1	17·99			
	3 3 0	18·00			
21	1 0 3	18·94	18·98	25	24·5
	4 1 1	18·99			
22	1 1 3	19·94	19·98	12	10·7
	3 3 1	19·99			
	4 2 0	20·00			
23	3 2 2	20·97	20·91	20	18·8
24	2 0 3	21·94	21·93	8	8·2
	4 2 1	21·99			
25	2 1 3	22·94	22·91	4	3·8
	4 0 2	23·97			0·2
26	4 1 2	24·97	24·94	2	0·7
	5 0 0	25·00	—	—	0
	3 4 0	25·00	—	—	0
27	2 2 3	25·94	25·93	25	18·7
	3 3 2	25·97			
	5 1 0	26·00			
28	3 0 3	26·94	26·96	3	3·5
	5 0 1	26·99			
	3 4 1	26·99			
29	3 1 3	27·94	28·01	20	16·1
	4 2 2	27·97			
	5 1 1	27·99			
	5 2 0	29	—	—	0
30	3 2 3	30·94	30·95	15	11·1
	5 2 1	30·99			
	0 0 4	31·90	31·98	3	2
	4 4 0	32·00			
31	1 0 4	32·90	33·00	25	16·3
	4 3 2	32·97			
	5 0 2	32·97			

TABLE II. (*continued*).

No.	$h_1 h_2 h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)^2$.	I obs.	I cal.
32	1 1 4	33·90			
	4 0 3	33·94			
33	5 1 2	33·97	33·92	25	12
	4 4 1	33·99			
	5 3 0	34·00			
34	4 1 3	34·94	34·94	5	3
35	2 0 4	35·90			
	3 3 3	35·94	36·01	25	19
36	5 3 1	35·99			
	6 0 0	36·00			
37	2 1 4	36·90			
	5 2 2	36·97	36·94	4	1·8
	6 1 0	37·00		—	0
38	4 2 3	37·94	38·00	30	22·8
	6 0 1	37·99			
	6 1 1	38·99			1·5
39	2 2 4	39·90			
	4 4 2	39·97	40·05	15	9·2
	6 2 0	40·00			
	3 0 4	40·90	40·85	7	4·4
40	3 1 4	41·90	41·93	18	12·6

TABLE III.—$N(CH_3)_4Br$.

No.	$h_1 h_2 h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)$.	I obs.	I cal.
	1 0 0	1			
1	0 0 1	1·96	1·94		
	1 1 0	2·00			
2	(1 0 1) β	(2·40)	(2·36)		
3	1 0 1	2·96	2·94	20	20
4	(1 1 1) β	(3·30)	(3·31)		
5	1 1 1	3·96	3·95	100	100
	2 0 0	4·00			
	2 1 0	5·00	—	—	0
6	(2 1 1) β	(5·75)	(5·75)		
7	2 0 1	5·96	5·93	5	5
8	2 1 1	6·96	6·94	60	49·2
9	(1 0 2) β	(7·31)	(7·32)		
10	0 0 2	7·85	7·953	7	9·4
	2 2 0	8·00			
11	(2 2 1) β	(8·23)	(8·27)		
12	1 0 2	8·85	8·832	45	26·7
	3 0 0	9·00	0	—	0
3	2 2 1	9·96	9·94	42	24·8
	3 1 0	10·00			
	3 0 1	10·96	—	—	0

TABLE III. (*continued*).

No.	$h_1\,h_2\,h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)^2$.	I obs.	I cal.
14	2 0 2	11·85	11·925	36	20·2
	3 1 1	11·96			
15	2 1 2	12·85	12·847	30	14
	3 2 0	13·00	—	—	0
16	(3 0 2) β	(13·93)	(13·97)		
17	3 2 1	14·963	14·934	10	7·3
18	2 2 2	15·85	15·90	—	1
	4 0 0	16·00		10	9·7
19	3 0 2	16·85	16·872	35	28
	4 1 0	17·00	—	—	0
20	0 0 3	17·667	17·974	—	1·3
	3 1 2	17·85		—	2
	4 0 1	17·903		15	11·8
	3 3 0	18·00			
21	1 0 3	18·67	18·896	28	4·4
	4 1 1	18·96			20·40
22	1 1 3	19·67	19·94	7	6·8
	3 3 1	19·96		12	8·9
	4 2 0	20·00 *d*			
23	3 2 2	20·85	20·87	32	24
24	2 0 3	21·67	21·81	10	4·3
	4 2 1	21·96			9·5
25	2 1 3	22·67	22·71	10	10·2
	4 0 2	23·85			0·1
26	4 1 2	24·85	24·85	10	6
	5 0 0	25·00	—	—	0
	3 4 0	25·00	—	—	0
27	2 2 3	25·67	25·79	10	9·8
	3 3 2	25·85		—	1·5
	5 1 0	26·00		—	1·9
28	3 0 3	26·67	27·06	—	2·2
	5 0 1	26·96		7	6·8
	3 4 1	26·96			
29	3 1 3	27·67	27·98	23	1·7
	4 2 2	27·85			1·4
	5 1 1	27·96			14·5
	5 2 0	29	—	—	0
30	3 2 3	30·67	30·71	14	5·2
	5 2 1	30·96			9·7
	0 0 4	31·41	—	—	1·9
	4 4 0	32·00	31·95	2	3·1
	1 0 4	32·41			
31	4 3 2	32·85	32·90	35	21·6
	5 0 2	32·85			
32	1 1 4	33·11	32·52	7	4·7
	4 0 3	33·67			3·3
33	5 1 2	33·85	34·00	7	3
	4 4 1	33·96			5·7
	5 3 0	34·00			

TABLE III. (*continued*).

No.	$h_1 h_2 h_3$.	$h_1^2+h_2^2+h_3^2\left(\frac{a}{c}\right)^2$.	$\left(\frac{2a}{\lambda}\sin\phi\right)^2$.	I obs.	I cal.
34	4 1 3	34·67	34·80	8	6·5
35	2 0 4	35·41	35·59	7	4·5
	3 3 3	35·67			
36	5 3 1	35·96	36·04	25	14·7
	6 0 0	36·00			1·0
37	2 1 4	36·41	36·88	10	
	5 2 2	36·85			6
	6 1 0	37·00			0
38	4 2 3	37·67	37·67	25	15
	6 0 1	37·96			
	6 1 1	38·96	38·90	4	3·2
	2 2 4	39·41	—	—	1·7
39	4 4 2	39·85	40·06	15	0·7
	6 2 0	40·00			8·7
	3 0 4	40·41			1·0
40	3 1 4	41·41	41·40	25	14·8

TABLE IV.—$N(CH_3)_4I$.

$h_1 h_2 h_3$.	I cal.	I cal.	$h_1 h_2 h_3$.	I cal.	I cal.
0 0 1	15	25	3 3 1	12	5·4
1 1 0			4 2 0		8·1
1 0 1	15	20·4	3 2 2	25	23·9
1 1 1	100	100	2 0 3	—	1·4
2 0 0			4 2 1	30	12
2 1 0	—	0	2 1 3		15·6
2 0 1	12	12·4	4 1 2	15	8·6
2 1 1	45	38·0	5 0 0		0
2 2 0	15	12·3	3 4 0		0
1 0 2	30	26·7	2 2 3		3·9
3 0 0	—	0	5 1 0	10	4·16
2 2 1	40	30·7	3 0 3		6·6
3 1 0			5 0 1	10	12·1
2 0 2	—	1·6	3 4 1		
3 1 1	30	27·3	4 2 2	15	3·68
2 1 2	30	20·7	5 1 1		15·0
3 2 0	—	0	5 2 0	—	0
3 2 1	6	6·8	3 2 3	10	11·38
2 2 2	6	2·6	5 2 1		6·36
4 0 0		9	0 0 4		1·61
3 0 2	25	22·7	4 4 0	—	2
4 1 0	—	0	1 0 4	—	1
4 0 1	22	4·83	4 3 2	30	12·94
3 3 0		9·71	5 0 2		8·35
1 0 3		9·55	1 1 4		5
4 1 1	15	16·6	4 0 3	—	1·5
1 1 3		2·7	4 4 1	22	4·33
			5 3 0		6·89
			4 1 3		10·14

We then proceeded to calculate the intensities of the lines on the basis of the atomic arrangement previously found * for the iodine compound.

Introducing the same angular parameters for all three compounds, we found the intensities given in (B) figs. 1 and 2. In fig. 3 (B) are given the corresponding intensities

Fig. 1.

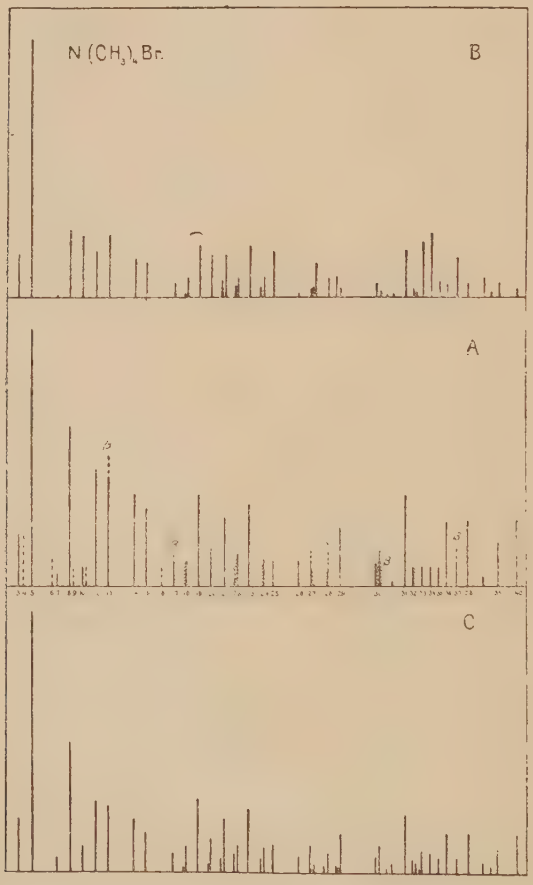
for the iodine compound. Comparing observed and calculated intensities we find that the agreement, which, as stated in a previous paper, is fairly good in the case of the I compound, becomes less good in the case of the Br compound, and still worse in the case of the Cl compound.

* *Loc. cit.* (II.).

After having tried various changes of the position of the atoms, we finally found that the following re-arrangement led to a satisfactory agreement between observed and calculated intensities.

In the lattice previously found the N-atoms were fixed with one parameter. With the particular value found for

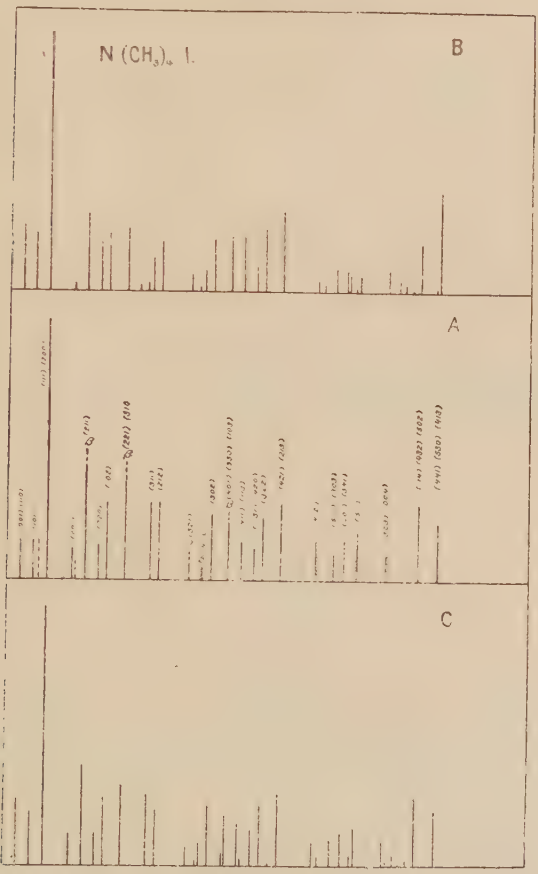
Fig. 2.



the parameter it appeared that the two nitrogen atoms contained in the unit cell were very nearly situated in a plane perpendicular to the *c*-axis. The C-atoms formed tetrahedral groups with their centres in level with those of the corresponding nitrogen atoms.

The re-arrangement of the atoms which had to be introduced mainly consisted in moving the nitrogen atoms a distance $a/2$ parallel to a, so that they were placed in the centres of the carbon tetrahedra. We now further assumed that the two nitrogen atoms contained in the unit cell were exactly placed in a plane perpendicular to the c-axis. The

Fig. 3.

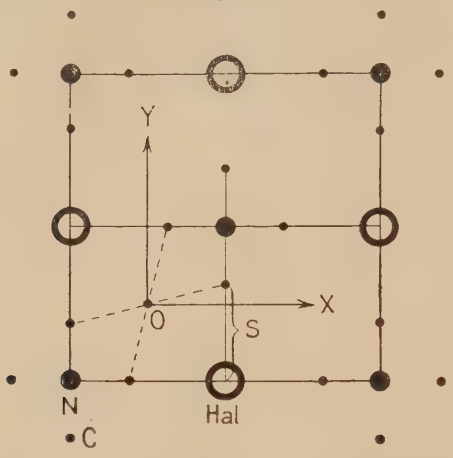


positions of the N-atoms are thus fixed without any parameter.

This change of position of the N-atoms does not involve any change of the space group of the lattice, and we may also say that the nitrogen atoms occupy the position of the space group D_{4h}^{7}, which has no parameter.

We now naturally take the plane of the nitrogen atoms as (X—Y) plane of the coordinate system.

Fig. 4.



Taking the symmetry centrum (O) midway between the two neighbouring N-atoms (fig. 4) as origin of the co-ordinates, the positions of the atoms may now be written in the form:

$$\left.\begin{array}{rl} \text{N}: & \pm\left[\left[\tfrac{1}{4},\ \ \tfrac{1}{4},\ o\right]\right] \\ \text{Cl, Br or I}: & \pm\left[\left[\tfrac{1}{4},\ -\tfrac{1}{4},\ p\right]\right] \\ \text{C}: & \pm\left\{\begin{array}{l}\left[\left[\tfrac{1}{4},\ (-\tfrac{1}{4}+s),\ -r\right]\right] \\ \left[\left[\tfrac{1}{4},\ (-\tfrac{1}{4}-s),\ -r\right]\right] \\ \left[\left[(\tfrac{1}{4}+s),\ -\tfrac{1}{4},\ -r\right]\right] \\ \left[\left[(\tfrac{1}{4}-s),\ -\tfrac{1}{4},\ -r\right]\right]\end{array}\right. \end{array}\right\} \quad . \ . \ . \quad (1)$$

Indicating the number of electrons associated with one of the halogen atoms by A and disregarding the reflexion power of the hydrogen atoms, the structure factor takes the simple form:

$$\left.\begin{array}{l} h_1+h_2=2n, \\ \qquad \text{S}=\text{N}(-1)^{h_1}+\text{A}\cos h_3\alpha+2\text{C}\cos h_3\gamma\left[\cos h_1\delta+\cos h_2\delta\right] \\ h_1+h_2=2n+1, \\ \qquad \text{S}=\qquad\quad -\text{A}\sin h_3\alpha+2\text{C}\sin h_3\gamma\left[\cos h_1\delta+\cos h_2\delta\right], \end{array}\right\} \quad . \ . \ . \quad (2)\ *$$

* After this result was obtained we found that P. Niggli (*Zeitschr. f. Kristall.* liii. p. 210, 1922), in discussing Vegard's original atomic arrangement, has suggested that possibly the N-atoms might take up that position of the space group which is fixed without parameters.

where

$$\alpha = 2\pi p,$$
$$\gamma = 2\pi r,$$
$$\delta = 2\pi s.$$

Comparing the parameters here introduced with those of the previous paper (II.), we have to remember that

$$p = 1/4 + p',$$
$$r = 1/4 - r',$$

where p' and r' are the corresponding parameters previously used.

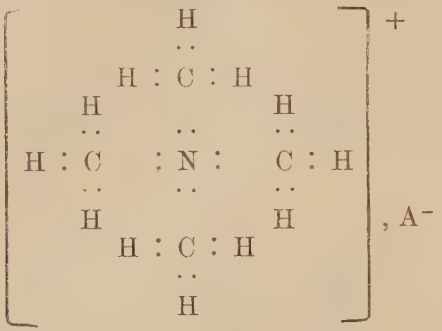
In order to find the typical intensity distribution, we calculate the intensities by means of the simplified formula :

$$J = \frac{\nu s^2}{h_1^2 + h_2^2 + \left(\frac{a}{c}\right)^2 h_3^2}. \quad . \quad . \quad . \quad . \quad (3)$$

The atomic arrangement previously found and described in the papers referred to (I. and II.) was most simply described by means of molecular elements $N(CH_3)_4I$, where the N- and I-atoms are situated on the same fourfold axis. When the N-atoms are put into the centres of the C-tetrahedra, and we get the arrangement here described, the molecular elements lose their significance, and we have a type of arrangement more like that of rock-salt. This fact would indicate that the crystals under consideration are forming lattices with polar binding (ionic lattices), where the halogen forms the negative ion, and the complex $N(CH_3)_4$ as a whole the positive ion.

In the formula for the structure factor we should put A equal to N + 1, where N is the atomic number of the halogen atom. With regard to the group $N(CH_3)_4^+$, the calculations show that it cannot be treated as a single centre with a weight equal to the sum of the electrons, but we have to regard each atom as a separate centre with its own electrons. An exact calculation of the reflexion power of the group would involve a knowledge of the way in which the electrons on an average are distributed within the complex, which we do not possess at the present time. Assuming a homœopolar binding of the atoms inside the complex, and assuming that the constitution formula may be written so as to express the tendency of the N- and C-atoms to surround themselves with groups of 8 electrons, and the tendency of the H-atoms to be attached by means of pairs (helium group), the constitution

formula of our substances might be written :



This constitution, however, as far as the "binding" electron systems is concerned, may merely have a formal significance. As a matter of fact, we found that a very good agreement between observed and calculated intensities is obtained by assuming the electrons to be associated with their respective nuclei, and it makes little difference for the calculated intensities whether we assume the halogen to be an ion or a neutral atom. Our final intensity values were calculated on the assumption that the reflecting power of the halogen is proportional to $(N+1)$ and that of C and N proportional to the respective atomic numbers.

In order to obtain approximate values of the parameters, we have also disregarded the reflexion power of the hydrogen atoms, for, as the reflexion power of each atom is very small, and as the H-atoms are spread out throughout the interior of the elementary cell, they should not very essentially affect the typical intensity distribution of the lines.

The parameter values of the halogen and C-atoms, which we found for the three compounds, are given in Table V.

TABLE V.

		$N(CH_3)_4Cl$.	$N(CH_3)_4Br$.	$N(CH_3)_4I$.
Halogen	p	0·361	0·375	0·394
	α	130°	135°	142°
Carbon	r	0·156	0·153	0·164
	γ	56° (56°·3)	56° (55°)	60° (59°·1)
	s	0·339	0·339	0·347
	δ	122°·5 (122°·1)	122°·5 (122°)	125°
Centre distance.				
N - C		1·48 Å.	1·49 Å.	1·51 Å.

The intensity distribution calculated in this way is given in Tables II., III., IV., and is graphically represented in (C) figs. 1, 2, and 3. We notice that the agreement between observed and calculated intensities is remarkably good for all three compounds.

The arrangement of the atoms N, C, and halogen corresponding to these parameters is illustrated in fig. 5 (Pl. XXI.), which shows a stereoscopic reproduction of a model of the lattice.

If we calculate the centre distances between the N- and the surrounding C-atoms, we find a value of about 1·5 Å., corresponding to the centre distance between neutral nitrogen and carbon atom in contact. We may thus regard the group NC_4 as forming a close packing of neutral atoms.

The Dimensions and Arrangement of the Hydrogen Atoms.

§ 3. Assuming the halogen to occupy the large space characteristic of univalent negative ions, we find that they have no direct contact with the NC_4-group. In order to explain the lattice as a packing of spheres, we must, as already found in our previous investigations, give a certain diameter to the hydrogen atom. As the N- and C-atoms are in contact, the hydrogen atoms must fill the space round the NC_4-group, and by means of "contacts" with the C-atoms and halogen ions, establish the stability of the lattice.

On account of the small reflexion power of hydrogen, it is hardly possible to determine exactly the position of the hydrogen atoms by means of the X-ray data alone, but as we may regard the atomic centres of N, C, and halogen as known, it might be possible to find the arrangement and dimensions of the hydrogen atoms from the assumption that the lattice may be regarded as a packing of spheres.

The elementary cell contains 24 hydrogen atoms. From the space group D_{4h}^7, we find that these atoms may either be divided into three groups of 8 atoms, or into two groups—one containing 8 and one containing 16 atoms. The first possibility does not seem to lead to any satisfactory arrangement, when all hydrogen atoms shall have equal diameters.

In order to fix the position and diameter of hydrogen, we must give definite diameters to N, C and the halogen ions.

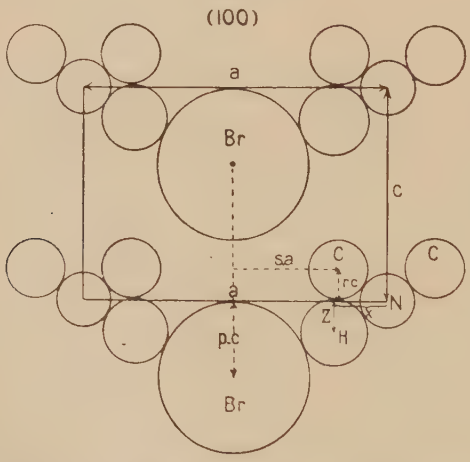
The values used are given in Table VI. In the case of N and C the diameters are supposed to be those of the neutral atoms, but as the centre distances N—C are not exactly equal for all three compounds, we have to introduce corresponding variations of the diameters.

The group of 8 atoms have their centres in (1 0 0) planes through the N-atoms. Their position is fixed by two parameters (x', z') giving the coordinates relative to an N-atom. These atoms may be placed in such a way that they have contact with both C and N and with the negative halogen ion. The position is illustrated in fig. 6, and from these three contact conditions the parameters x, z and the radius can be found.

The results, which were found by means of a graphical method, are given in Table VI.

In the group of 16 hydrogen atoms, each atom has three parameters (x, y, z) (referred to the centre of an N-atom).

Fig. 6.



By studying the space to be reserved for these atoms we found that each atom might have contact with two other H-atoms, one C-atom, and one halogen atom. This makes in all four contact conditions, which lead to the following equations :—

$$\text{I.} \quad [x-a(\tfrac{1}{2}-s)]^2+y^2+(z-cr)^2 = (r_c+r_H)^2,$$

$$\text{II.} \quad (\tfrac{1}{2}a-x)^2+y^2+[(1-p)c-z]^2 = (r_{Hal}+r_H)^2,$$

$$\text{III.} \quad 2(\tfrac{1}{2}a-x-y)^2 = (2r_H)^2,$$

$$\text{IV.} \quad 2(x-y)^2+[(1-2r)c-2z]^2 = (2r_H)^2.$$

From these four equations we should be able to find the four unknowns $x, y, z,$ and r_H. A direct solution of these equations would be somewhat troublesome, and we have

used a method with trial and error. From the model of the space lattice we can easily find certain limits for the values of r_H. We then introduce a probable value of r_H, calculate $(x\ y\ z)$ by three of the equations, and use the fourth as a test. In this way we found the parameters and dimensions given in Table VI.

A stereoscopic reproduction of a model showing the arrangement of the H-atoms relative to the NC_4-group and the halogen is given in fig. 7 (Pl. XXI.).

TABLE VI.*

		$N(CH_3)_4Cl$.	$N(CH_3)_4Br$.	$N(CH_3)_4I$.
$r_{Hal.}$ (ion)		1·82 Å.	1·96 A.	2·15 Å.
r_N		0·71 „	0·71 „	0·72 „
r_C		0·77 „	0·78 „	0·79 „
Groups of 8 H-atoms.	x/a	0·184	0·183	0·177
	z/c	0·140	0·182	0·122
	r_H	0·86 Å.	0·85 Å.	0·85 Å.
Groups of 16 H-atoms.	x/a	0·224	0·215	0·197
	y/a	0·117	0·127	0·180
	z/c	0·386	0·373	0·356
	r_H	0·85 Å.	0·85 Å.	0·87 Å.

We notice the remarkable fact that the *radius of the* H-*atoms comes out nearly equal for both groups of* H-*atoms and for all three compounds and equal to* 0·85 Å.

The mere fact that the centre distance N – C, deduced from X-ray data, corresponds to that of contact between these two atoms, shows—as already mentioned—that the hydrogen atoms must require a space for themselves.

In order to verify this conclusion we have calculated the intensities on the assumption that CH_3 is forming spherical groups in contact with N and halogen, but the result was not in agreement with observations. If we shall regard the structure as a packing of spheres, *the hydrogen atom must be given the fairly large diameter of about* 1·70 Å.

If we would drop the idea of a radius for the H-atoms and only consider the dimensions of the group CH_3, this group could not be given a spherical shape if stability is to be secured by mutual contacts, but, as stated by one of us in a

* By the calculation of the numbers given in Table VI. we have used values of the parameters $(\gamma\ \delta)$ of the C-atoms, which are given in the brackets of Table V.

previous paper (II.), the CH_3-group should be given some other shape, *e. g.* that of a rotational body like an egg.

Although we cannot claim any great accuracy as regards the parameters of the hydrogen atoms, still the fact that we have obtained a nearly constant value for the hydrogen diameter indicates that they cannot be far from the true values.

If so, the present substances should give us a means of fixing more accurately how the electrons of the complex $N(CH_3)_4^+$ are distributed among the atoms. The procedure would be to calculate the intensities for a number of different possible distributions and compare them with observations.

It would especially be of interest to find out whether the hydrogen atom reflects with a power corresponding to *one* or to *two* electrons associated with it. In the latter case hydrogen might be regarded as an univalent negative ion with an electron system like that of the neutral helium atom.

Such an investigation would lead to a considerable calculating work and will be left for a later publication.

The morphotropic relation between the substitution products $N(CH_3)_4$ hal. and their mother substances NH_4 hal. was studied by one of us in the paper of 1917 (I.). This relation is not at all so simple as the comparison between the topic parameters would suggest. The number of molecules in the unit cell is different for the mother substance (4) and for the substitution product (2). Comparing the dimensions of the unit cells there is apparently no simple connexion. If, however, we determine the dimensions (a', c') of a similarly shaped cell containing only one molecule and compare it with the side of a cube containing one molecule of the mother substance, a simple relation comes out, which is apparent from Table VII.

Table VII.

	a'.	c'.		a'_M.
$N(CH_3)_4I$	6·278	4·556	NH_4I	4·54
$N(CH_3)_4Br$	6·118	4·366	NH_4Br	4·328
$N(CH_3)_4Cl$	6·023	4·265	NH_4Cl	4·143

The side of the cube of one molecule for the mother substance (a'_M) *is approximately equal to the height* (c') *of the prism containing one molecule of the substitution product.*

In this connexion it is of interest to notice that the ammonium salts can appear in a form of the (CsCl) type,

where the unit cell contains one molecule, but the sides of these cubes for the same substance are smaller than a'. They were found * to be 3·88 Å. and 4·07 Å. for the Cl and Br compounds respectively.

Thus we see that the volume occupied by one molecule very much depends on the atomic arrangement, and it is then not easily understood why the dimension in the *c*-direction of the one-molecular cell should be practically unchanged by a substitution which essentially alters the atomic arrangement.

Summary of Results.

1. Investigations undertaken in 1917 on the structure of $N(CH_3)_4I$ have been continued by using the powder method and extended to the bromine and chlorine compounds.

2. The correctness of the dimensions of the unit cell and of the space group (D_{4h}^7) originally found for the structure of $N(CH_3)_4I$ has been confirmed.

3. The investigation of the bromine and chlorine compound, where the reflecting power of the halogen atoms is less dominating than in the case of the iodine compound, has made it possible to fix more accurately the position of the nitrogen and carbon atoms.

4. The nitrogen and carbon atoms form groups NC_4, where N is placed in the centre of a tetrahedron of C-atoms. The centre distance N—C of these groups is found to be 1·50 Å., corresponding to that of neutral atoms.

5. If we regard the crystal as a packing of spheres, the hydrogen atoms must be given a definite diameter. In order to obtain the same diameter for all hydrogen atoms in the crystal we have to assume one group of 8 and one group of 16 H-atoms inside the unit cell. From the volume conditions we can calculate the parameters and diameter for each group of H-atoms. Both groups give the same value, 1·70 Å., for the diameter of the H-atom.

6, The determination of the positions of the H-atoms from volume conditions may give us a possibility of determining the number of electrons connected with each H-atom in the space lattice of the crystals.

7. The arrangement of the atoms suggests that the crystal has an ionic constitution $N(CH_3)_4^+ - Hal$.

Physical Institut, Oslo.
July 4, 1927.

* L. Vegard, *Zeitschr. f. Phys.* v. p. 21 (1921).

XCIII. *Condenser Discharges in Discharge-Tubes.*—Part I. *Single Condenser Discharges. By* William Clarkson, *M.Sc., A.Inst.P., The Physical Institute of the University of Utrecht* *.

1. Introduction.

FOLLOWING a sequence of observations on the volt-ampere characteristics in discharges [1], this paper applies the observed phenomena to condenser and intermittent discharges through gases to explain their critical voltage relations.

Each condenser discharge is considered as presenting the following sequence of phenomena:—The "striking" of the discharge, the consequent "build-up" of current, the "extinction," and the "clear-up" of space-charge.

It has been demonstrated recently that the potential at which a discharge "strikes" is variable according to the current and space-charge in the tube. The idea of a "threshold current" was utilized by the author to explain the progressive lowering of the "striking" potential with increase in frequency in intermittence [1]. It has been shown that the "threshold-current" characteristic is identical with the corona characteristic [2]. This identity is assumed here also for dynamic discharge conditions, since this simplifies presentation.

It was well known that a "lag" could take place at the "striking" of a discharge; but whether it occurred before or after this initiation of a self-sustained discharge has been much discussed [3]. It was recently shown that a "lag" occurred even in very clean discharge-tubes. The attribution of this "lag" to a slow "build-up" permitted the peak-voltage phenomena in "intermittence" to be explained [4].

The "clear-up" of both steady and condenser discharges has been studied in detail here, for the variation of the minimum potential in both single "flashes" [5] and "intermittence" [6] has been attributed to this phenomenon.

2. Tubes and Apparatus.

The type of tube used here has been described previously [4]. The electrodes were of nickel, convex, about 5 cm. in diameter and 1 cm. apart. The tubes were thoroughly cleaned and baked-out before filling with purified gas: after filling, the electrodes were sodiumated electrolytically.

Tubes containing neon at pressures of 4 mm. and of

* Communicated by Prof. Ornstein.

8 mm., and with argon at 10 mm., were employed. The discharge phenomena observed were alike in all cases, and did not differ essentially from the phenomena observed on previous occasions.

The measuring apparatus has also been described. A quick-acting electrostatic voltmeter was used for voltage measurements, and a shunted galvanometer for the currents. The circuit current was controlled by a diode valve.

3. Condenser-Discharge Characteristics.

a. *The Volt-ampere Characteristic.*—The characteristics found were of the usual form, as shown in fig. 1 (*a*), where SR is the corona characteristic, RQP the "stable" characteristic, and the stage beyond Q the "normal" characteristic.

The currents corresponding to the corona characteristic were the minimum (i_n), for the applied voltage (v_n), at which a self-sustained discharge was possible. Space-charge was a factor of great importance here. v_n was constant, at the value of the sparking potential (v_c), for currents of a few microamperes, but then fell rapidly as i_n increased.

The corona characteristic has also been referred to as the "threshold-current" characteristic, since, under stable conditions at least [2], it determines the potential at which a discharge "strikes" and at which a "build-up" is possible (v_m, i_m).

The normal characteristic gives the maximum current sustainable at the applied voltage.

We may assume that on "build-up" the current increases, from a point v_m, i_m, until another point on the characteristic, v_n, i_n, is realized. Extinction will then occur [1].

Experiments show that on the "extinction" of a steady discharge the voltage falls by an amount variable with conditions. This is referred to as the "clear-up" of space-charge.

Referring to a diagram of the characteristic, with voltage as ordinate and current as abscissa, we see that "build-up" must occur in the region to the right of the characteristic, and "clear-up" in the region to the left. These are regions of instability, and only realizable dynamically (see later).

Considering that multiplication of ions is possible for all except small voltages, the characteristic is definable as the boundary where the factors tending to increase the number of ions are exactly counterbalanced by those tending to reduce the number. To the right the former factors predominate, to the left the latter.

We may assume that all points to the right of the

characteristic are realizable; but we find that the same assumption applied to the region to the left, would imply that currents at v_n greater than the corresponding i_n were possible for points on the normal characteristic. The normal characteristic may give the maximum current sustainable with an abnormal cathode fall only, in which case we may suppose that greater currents may be attainable under other space-charge conditions. Though the attempts made to realize these conditions by photo-electric means proved inconclusive, the P.E. characteristics given by Campbell[7] offer support to the idea. Certainly, if only normal cathode falls were possible, v_n would remain constant at v_b, and this problem would not arise.

Intermittence and corresponding experiments, and the fact that the "clear-up" associated with the characteristic increases regularly with current, suggest that the normal characteristic is not necessarily traversed towards Q from v_n, i_n in flashes. Below Q, of course, "extinction" must occur, since the voltage, necessarily decreasing, is already a minimum for the self-sustained discharge.

b. V, I *Diagrams.*—Fig. 1 (*a*) shows the current-voltage relations assumed in this paper for single condenser discharges. From the striking of the discharge at (nominally) zero current, a voltage $V_M \Rightarrow v_C$ being necessary, the current increases steadily, with a corresponding decrease of the condenser potential, until a voltage, v_n, determined by the current and the space-charge distribution at which the discharge is no longer self-sustainable, is reached. "Extinction" occurs here and the current falls to zero, the condenser voltage being meanwhile lowered to a final value V_N (the "tracks" drawn in fig. 1 (*a*) are diagrammatic, not experimental).

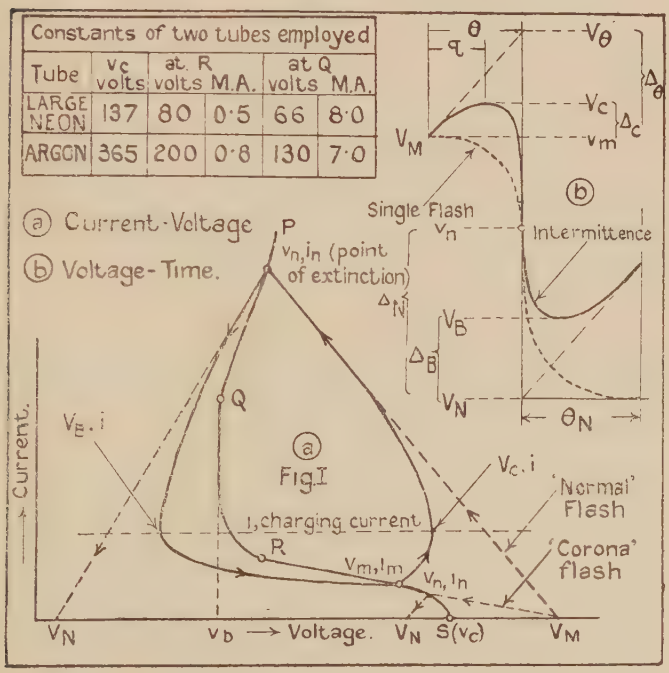
In addition to the assumption that "extinction" occurs at v_n, i_n, as would be determined by the characteristic under statie conditions, it is further assumed that the "clear-up" at the characteristic is representative of that of "flashes" also. Both these are necessary experimental assumptions. They probably give results not only qualitatively correct, but quantitatively of the right order.

One point must be stressed at this stage. To define any stage in a discharge accurately, in addition to the values of voltage and current, as given in the diagram, the space-charge distribution must also be known. This is indeterminable, and is variable. It must be connected with the

time duration of the discharge, another factor that cannot be included in such a diagram, though since this is measurable we are not without some insight into the conditions actually obtaining during the discharge.

Only points on the characteristic, for steady discharges, where the space-charge conditions are reproducible, are therefore definite. All others are indeterminate; and in

Fig. 1.



condenser discharges, particularly in intermittence, it may be said that no points are fully definable. It is for this reason that the assumptions as to the significance of the (statical) characteristic in discharges, in "striking" and "extinction," are necessary. They are made with full cognizance of the above facts. It is doubtful whether we may speak of an unvarying characteristic in discharges at all; but, in any case, the usual volt-ampere characteristic presents the only concrete approximation that may be utilized.

4. "Clear-up" at Characteristic.

The clear-up of space-charge presents many points of intrinsic interest, and the writer decided to attempt as thorough a quantitative examination as was possible.

a. *Method.*—A condenser, shunted by the electrostatic voltmeter, was placed across the tube, through which a steady current was then passed. Some point (v_n, i_n) on the characteristic was thus attained. When the current circuit was suddenly broken the discharge ceased, but not before a quantity of electricity, Q, sufficient to lower the condenser voltage to V_N, $Q = C(v_n - V_N)$, had been transferred. Given constant conditions it was found that V_N had a definite value for any one current and capacity.

It was obviously not possible to study the whole of the characteristic in this way, as the range over which steady currents were possible depended on the capacity; but from qualitative results, obtained with a diode in the condenser circuit (intermittence thus being prevented), it was seen that the conclusions drawn from the original method were of general applicability.

b. *Results.*—All these tubes showed the same general clear-up properties. $\Delta_N (= v_n - V_N)$, increased with increase of i_n, and decreased with increase of C. As far as could be determined, the corona characteristic and stable characteristic gave a continuous Δ_N, i curve.

Fig. 2 (*b*) shows the variation of Δ_N (large neon tube) for capacities from 0·0022 μF.–0·9 μF., and for currents (roughly) from 0·1–50·0 M.A., i_n being plotted on a log scale. It will be seen that the rate of increase of Δ_N falls off with increase of current, and that V_N approaches asymptotically a small limiting value. This is more easily demonstrable with small capacities, the currents required for larger capacities being beyond the practical range, being about two amperes for a capacity of 1·0 μF.

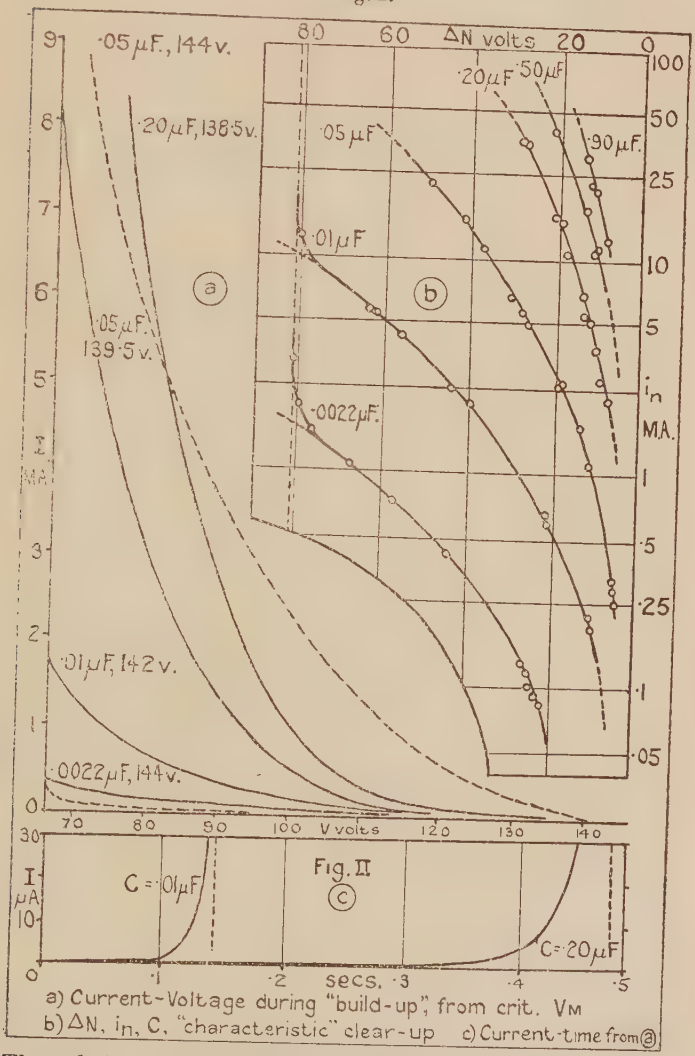
As the variations of v_n with i_n were small enough to be neglected, fig. 2 (*b*) actually gives the Δ_N, C, i_n variations for a constant initial voltage. Their relation is given accurately by the empirical expression

$$\Delta_N = k \frac{i_n^x}{C^y} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

In this case x and y were approximately of the same magnitude (0·5 and 0·55 respectively), and k had a value of 0·07.

It was found that the same expression applied to similar clear-up curves for the other tubes. The argon-filled tube, for instance, gave $k=0{\cdot}19$, and both x and y equalled $0{\cdot}42$.

Fig. 2.



Though it cannot definitely be stated at this stage that this relation is common to all discharge-tubes, whatever the electrode disposition etc., certain definite statements concerning clear-up may be made.

Since in the four tubes examined x and y always had values of the order of 0·5, it is proposed to simplify treatment by assuming this value throughout. Thus we may say from equation 1 that Δ_N varied as $\sqrt{i_n/C}$.

The quantity transference is given from equation (1) by

$$Q = k i_n^x C^{1-y}, \qquad \text{approx.} \propto \sqrt{i_n C}. \quad . \quad . \quad . \quad (2)$$

The fact that Q varies with C, i constant, indicates that it is not a constant quantity determined by the initial current, but that it also varies with voltage. Multiplication of ions during clear-up is thus involved, and these effects must be considered.

c. *General Considerations.*—Visual observation of the phenomenon of clear-up in intermittence shows that the positive column disappears almost immediately [6] [1]. Practically all the clear-up period is occupied by the diffusion of the negative glow towards the anode. It is possible that this period is characterized by currents of the order of corona characteristic currents. Whether the current experiences a discontinuity or sudden decrease at the moment of extinction, however, cannot be inferred.

Considerations previously applied to "build-up," [4] are extended to clear-up in the following treatment. Space-charge distribution is perforce neglected.

During any time in clear-up there will occur multiplication, recombination, and diffusion of ions. We may assume that the rate of change of space-charge falls rapidly with reduction of V. Current-voltage characteristics obtained for photo-electric currents (constant illumination) in the clear-up region, showed that I increased continuously with V. The rate of increase became greater as V approached v_b, a change of, say, 20 volts in this region bringing about a twentyfold change of current.

The rate of clear-up will be approximately constant at constant voltage.

If the current is assumed to depend on the amount of space-charge present in the inter-electrode space, a somewhat exponential time-current curve will result.

For the case where there is a capacity across the tube, the voltage during clear-up will fall at a rate determined by the current, most rapidly at first when I is a maximum, and then much more slowly, as the ensuing reduction in potential will have caused a correspondingly rapid diminution of current. The space-charge will persist longer at higher potentials.

The variation of Δ_N with C and i_n may now be considered.

At a constant rate of clear-up (i_n constant, and therefore Q constant), Δ_N should vary inversely as C, but as the rate of clear-up increases rapidly with diminution of V (*i.e.* with reduction of C), Δ_N will actually vary much more slowly than i/C. The same argument applies to changes with i_n (C constant), the change in rate of clear-up causing Δ_N to vary much less rapidly than the current i_n. We may further expect that the result of corresponding changes in i_n and C will be of the same order; and this, indeed, is the case, though the significance of the square-root relation ($\Delta_N \propto \sqrt{i_n/C}$) lies in the quantitative relations.

The relation between Δ_N and the voltage during clear-up was clearly demonstrated in V_N, i_n curves in cases where v_n changed rapidly with i_n (above Q, fig. 1, *a*). The curves showed a pronounced change in slope at Q, Δ_N increasing rapidly. The constants x and y were also found to change.

The variation of Q under different conditions is contained in the foregoing discussion.

d. *Times.*—It has been shown that with the voltage across the electrode zero, clear-up occurs in a time less than 10^{-6} second [8]; and since with V near v_b the time is large (∞ at v_b), it appears that a wide variation of time of clear-up with voltage is possible.

An idea of the relative variation of the time of clear-up may be obtained from a consideration of the quantity transference and the ratio of the average to the initial current. Quantitative results are impossible, since the actual value of the average current, though probably of the same order of magnitude as corona currents, is not known.

One effect of a lowering of V_N will be to reduce the relative duration of the final (the small current) stage in clear-up. The ratio initial current to average current will thus decrease as Δ_N increases, the actual value of the average current depending of course on i_n.

If we express this ratio by R, we have from equation (2) that

$$T \propto R\sqrt{C/i_n}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

or from equation (1) that T varies inversely as Δ_N, for R varies with Δ_N also.

The previous considerations are in agreement with this. We may expect that T (i_n constant, C variable) will increase as C increases, but with C constant and i_n variable will decrease, since Q varies much less slowly than i_n, and the average current varies (approximately) as i_n itself. We

may expect T to be (approx.) constant for constant V_N, since the relations during the clear-up will be the same, provided that i_n and C vary similarly, as in this case.

In the absence of a direct method of measuring these times, quantitative values were obtained for intermittence experiments. For the large neon-filled tube times of the order of 0·010–0·020 second were calculated for $\Delta_N = 40$–50 volts. If we accept these results, the limits of time in fig. 2 (*b*) are about 0·005–0·05 second. These are reasonable values. R is of the order of 30.

5. Single-Condenser Discharges.

a. *Method.*—The variation of the final voltage (V_N) with variation of the capacity (C), and the initial voltage (V_M), for single flashes, was examined by having the condenser, shunted by the voltmeter, arranged with a suitable key so that it could first be charged to a certain potential and then discharged through the tube. Capacities ranging from 0·002–1·0 μ F. were employed.

b. *Results.*—Results similar to those recorded in the previous work on the subject[5] were obtained for all tubes employed: a fully-developed (normal) discharge did not occur until V_M was definitely, often considerably, higher than v_c; V_N decreased with increase of Δ_M ($= V_M - v_C$), and the rate of change of V_N with Δ_M varied in the same manner. It was found, however, that this did not affect the validity of the definition of v_c as the "sparking potential," for observations in the dark showed that some form of charge always occurred with $V_M \gtrless v_c$; below the critical V_M they merely were of the corona form.

The full sequence of phenomena as presented most clearly, here with C of the order of 0·01 μ F., is described below.

At values of V_M just greater than v_c the discharge was scarcely visible. It consisted of a faint corona glow which welled out from the anode only after an appreciable lag. This lag became less and less in evidence as Δ_M increased, and the discharges became progressively brighter and of shorter duration. Though at first scarcely distinguishable from v_c, V_N decreased more and more rapidly as the critical V_M was approached. At a voltage just less than this critical voltage a bright corona flash of short duration was obtained, but at a voltage just greater the flash was normal and the

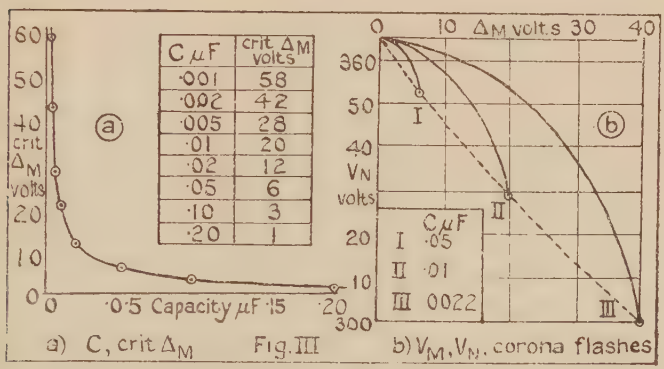
value of V_N fell considerably. It was often clearly seen, however, that the initial stages of these flashes were still corona glows.

In the previous work on the subject, V_N variations were explained from the various quantity transferences to be assumed as occurring for various capacities under different voltage and current conditions. With the assumptions made here as to the "extinction" and "clear-up" in flashes, these considerations assume a more precise significance. Further, actual values of i_n may be found from clear-up graphs when C and V_N are known.

Corona Flashes.

Fig. 3 shows the V_M, V_N variations (argon-filled tube) for various values of capacity. The curves are of the form just described for $C=0{\cdot}01\ \mu F.$, only the actual range of

Fig. 3.



variations changing with C (fig. 3, *b*). The critical V_M varied in a hyperbolic manner with C (fig. 3, *a*), the V_M, C curve being asymptotic to v_c. Δ_M was about 20 volts with $C=0{\cdot}001\ \mu F.$ V_M, V_N curves for other capacities could be constructed from these, since the form of the curve, the critical V_M, and the limiting value of V_N, were known in each case.

The curves show that with C constant i_n increased with Δ_M, since in all cases a lower V_N was associated with a higher V_M. Also that for V_M constant i_n increased with C, V_N here decreasing. The rapid changes of V_N in corona flashes are attributable primarily to the large decrease of voltage associated with increase of current, normal flashes

showing much less change, as v^n was approximately constant and only clear-up effects could vary.

These results showed that considerable lags were present at the initiation of discharges. Only when the current increased very slowly could the voltage be diminished sufficiently for extinction to occur on the early stages of the corona characteristic. Visual observation of the fainter flashes showed that lags of the order of 0·1–1·0 second occurred before the glow became bright. The lag decreased rapidly with increase of Δ_M.

Table I. gives the minimum value of θ, the time of duration of the discharge, for a wide range of values of Δ_M, three capacities being employed. θ was calculated from the quantity transference $C(V_M - v_n)$ and the maximum current i_n. V_N being known, i_n was found from the corona characteristic of the tube, due allowance being made for the depreciation of voltage due to clear-up.

Table I.

Capacity μ F.....	·10	·02			·0022				
Δ_M volts	·3	3·5	4·5	8·0	9·5	14	23	32	40
Min. $\theta \times 10^{-2}$ sec.	1·7	1·4	1·1	·45	·44	·18	·13	·09	·08

The Δ_M, θ curves were continuous and of hyperbolic form. The times calculated were all of appreciable magnitude. This is in excellent qualitative agreement with previous results. Quantitative comparison would demand that the average current was obtained; but that cannot be done with the accuracy required.

The hyperbolic form of the C, crit. V_M graphs (fig. 3, *a*) may be explained from these results. The time for v_c to be reached decreased with increase of Δ_M, but the current attained in this time increased. It is thus probable that the quantity transferred in this time will be at least approximately constant for a certain range of capacities.

"Normal" Flashes.

a. V_N, i_n, C, *relations.*—Experiments showed that in normal flashes, C constant, V_N increased with increase of Δ_M, as in corona flashes, and with Δ_M constant increased regularly with C, with the exception of the smallest capacities.

A minimum value of V_N of the order of 15 volts was recorded, as in the "characteristic clear-up." The maximum value of V_N was near to, though here always less than, v_b.

These V_M, V_N curves showed that i_n increased with, though less slowly than, C. With i_n directly proportional to C, V_N would be almost constant. The corresponding curves with C constant showed that i_n also increased continuously with C. Table II. illustrates the results obtained with the large neon-filled tube. They are in good agreement with the observed extent of the negative glow.

TABLE II.—Variation of i_n (M.A.) with C and Δ_M.

Δ_M volts.	Capacity μ F.			
	·0022	·01	·05	·20
10	·34	2·2	12	55
30	·61	3·8	22	79
50	·97	5·6	34	99

The currents associated with high values of C and Δ_M explain the occurrence of sparks in these cases. In connexion with sparking, it may be mentioned that the value of V_N did not immediately change to zero, but showed a progressive lowering with increase of C or V_M.

The table also clearly demonstrates the presence of lags, as only limited currents were attained in most cases.

b. V, I *Time, relations.*—These relations were obtained from the value of i_n at the critical V_M.

At a voltage just less than the critical V_M the discharge ends on the corona characteristic, and at a voltage just higher, on the normal characteristic. Consequently the current-voltage relations in this latter discharge should be obtained by joining V_M, 0, and v_a, i_n, by a curve just grazing the corona characteristic. The precise form of the curve, however, is not implied.

The time of discharge, θ, may be obtained from such by integrating 1/I, V graphs. In this case it is permissible to join the given points by a smooth curve, since I probably increases in a continuous manner. Instead of V, 0 being used however, it is obviously necessary to assume that I has some small initial value. The term "Zero" is permissible in descriptive treatment only.

The times calculated in this way with an initial current of 0·001–0·0001 μA., for capacities ranging from 0·002–0·20 μF., showed a variation of θ with V_M of precisely similar form to that previously obtained. The construction employed would thus appear to be justified, even though the actual values obtained were generally larger (up to $\times 2$) than the probably more correct ones.

Fig. 2 gives the V, I, curves implied, for all but the smallest currents; actually the voltage was almost constant until currents of the order of 0·01 μA. were attained. As is to be expected, in the presence of appreciable lags the greater part of the voltage-fall occurs before large currents are reached.

Fig. 2 (*c*) gives representative current-time curves. The presence of a "slow build-up" is strikingly demonstrated, the relative time to attain 10 μA. being from 0·85–0·95 θ. The corresponding voltage fall with the smallest capacity was 20 volts. Calculations of the ratio of the maximum to the average current gave 100–1000 as representative values. No direct relation is, of course, to be expected, since the ratio depends on the lag, and is only indirectly connected with the final current.

Fig. 2 (*a*) shows how the "rate of build-up" increases with voltage.

It must be pointed out that though these results do not preclude the existence of a short initial "time" lag [3], no direct evidence of such a lag has appeared up to the present.

By adopting the appropriate values for θ, it was possible to construct the I, V curves for values of V_M other than the critical one. The resultant curves differed from the original ones only in showing a much more rapid build-up (fig. 2, *a*).

The construction used above determines the relations of corona as well as normal flashes at the critical V_M, the only difference being that in the former case only the data up to the moment that the corona characteristic is reached are to be taken. Qualitatively the results agreed with those of Table I. As suggested, an appreciable curtailment of θ was apparent. Though the form of the curve has been obtained, nothing is to be gained by applying the ideas gained to all the corona flashes, though mention may be made that, as the ratios of the maximum to the average currents obtained here were of the order of 4–10, the time-results could with confidence be multiplied by some such (variable) factor. Such a treatment would not invalidate any of the conclusions drawn. It gives results in much better agreement with observed values.

Summary.

1. The role of the volt-ampere characteristic and the sequence of phenomena in the condenser discharges in discharge-tubes are considered.

2. The fall of voltage (Δ_N) during "clear-up" associated with various points on the characteristic (v_n, i_n), is studied with different capacities (C). Δ_N is found to be proportional to $\sqrt{i_n/C}$, and the time of clear-up, θ_N, to $1/\Delta_N$.

3. Such relations are qualitatively explainable from a rapid increase of clear-up with reduction of voltage.

4. The variation of the final voltage (V_N) with the capacity (C) and the initial voltage (V_M), corona and normal flashes, and the transition V_M, are studied. The maximum currents (i_n) attained are deduced from V_N and clear-up curves. i_n increases with C and V. Numerical values are given.

5. Corona flashes demonstrate a slow build-up. Their times of duration (θ) are obtained qualitatively.

6. The data at the transition V_M is employed to determine the actual current-voltage relations during discharge, and also to determine the duration of build-up (θ).

7. Current-time curves are of somewhat exponential form with a very slow increase of current initially. The rate of build-up increases rapidly with voltage.

References.

(1) Clarkson, Phil. Mag. (pending).
(2) Taylor, Phil. Mag. iii. p. 368 (1927).
(3) Zelany, Zuber, Peek, and others. See reference (4).
(4) Clarkson, Phil. Mag. iv. p 121 (1927).
(5) Taylor & Stephenson, Phil. Mag. xlix. p. 1081 (1925).
(6) Penning, *Phys. Zeit.* xxvii. p. 187 (1926).
(7) Campbell, Phil. Mag. iii. p. 925 (1927).
(8) Oschwald & Tarrant, Proc. Phys. Soc. xxxvi. p. 262 (1924).

XCIV. *Notices respecting New Books.*

La Théorie de la Relativité: Tome II. *La Relativité Générale de la Théorie de la Gravitation d'Einstein.* Par M. von Laue, Professeur de Physique Théoretique à l'Université de Berlin. Traduction faite d'après la quatrième édition allemande, revue et augmentée par l'auteur, par Gustav Létang. Pp. xvi + 318. (Paris: Gauthier-Villars et Cie. 1926. Price 78 francs.)

THE treatise of Dr. von Laue on the theory of relativity is one of the best that has been written, and the appearance of a French translation, assuring it a wider circulation, is to be welcomed.

Written by a physicist primarily for physicists, special care has been taken to smooth the difficulties of those who approach the theory for the first time, and to remove objections against the theory on the part of some who have failed properly to understand it, owing to an insufficient knowledge of non-Euclidean geometry and tensor calculus. A large portion of the volume is therefore occupied with matter introductory to the theory proper: an account is given of earlier theories of gravitation and of the physical ideas at the basis of the generalised principle of relativity. Two chapters deal with tensor calculus and non-Euclidean geometry. The deduction of the fundamental laws of physics, including electromagnetism, dynamics, and gravitation, is followed by the application of these laws to particular cases. The final chapters deal with the rigorous solution of the equations of the field, and with special developments of the theory. The physical insight of the author will appeal particularly to those who approach the theory from the physical rather than from the mathematical point of view.

Microskopische Physiographie der petrographisch wichtigen Mineralien. Begründet von H. Rosenbroch. Bd. I. *Erste Hälfte-Untersuchungsmethoden. Fünfte, völlig umgestaltese Auflage*, von Dr. E. A. Wülfing. Pp. xxiv + 848, with 680 text-figures and 15 plates. (Stuttgart: E. Schweizerbarische Verlagsbuchhandlung. 1921–24. n.p.)

The new edition of the great classical work by Rosenbuch on microscopical petrography contains nearly 400 more pages than the previous edition, which appeared in 1904. The first edition, published in 1873, contained only 111 pages, and each succeeding edition has shown an increase both in the number of pages and in the number of text-figures. This indicates the rapid development of the subject. It contains a comprehensive account of optical theory, including both geometrical and physical optics with special attention to crystal optics. The general theory of the microscope receives very thorough treatment, and many different types of petrological microscopes are described, as well as various accessories. A full account is given of various experimental methods with numerous practical details.

The present edition was issued in three portions, which appeared at intervals. It is now available in collected form. On account of its completeness and accuracy, it is a manual which no worker in the field of microscopical petrography can afford to be without.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

Fig. 2

Above: Air bubble rising in "Water-glass."
Below: Negative showing movement of bubble and liquid.

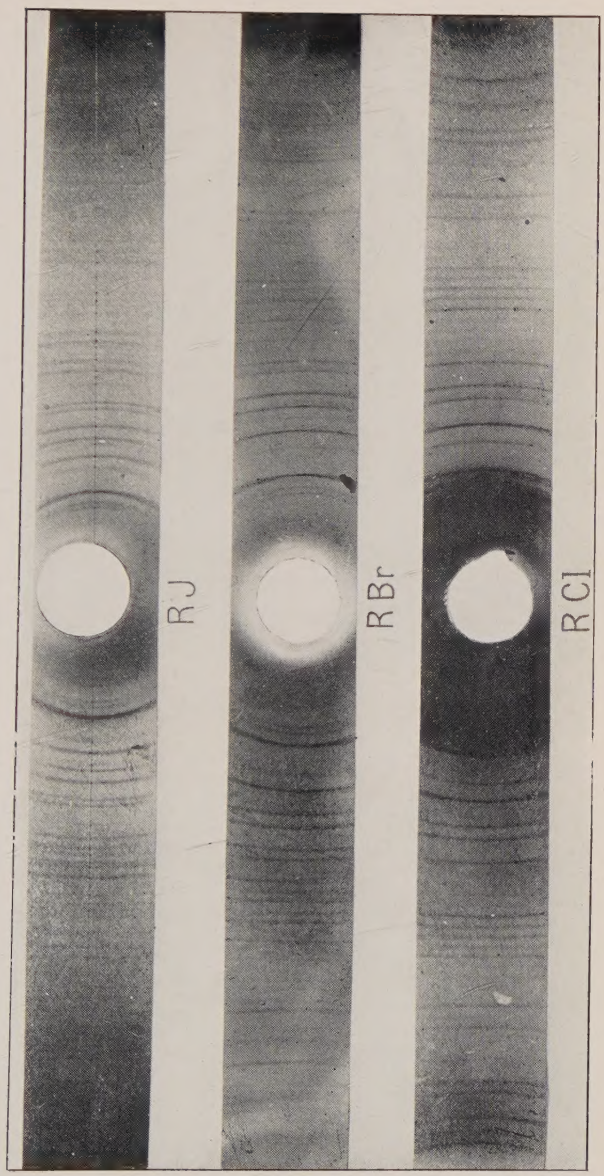

RJ
RBr
RCl

Fig. 5.

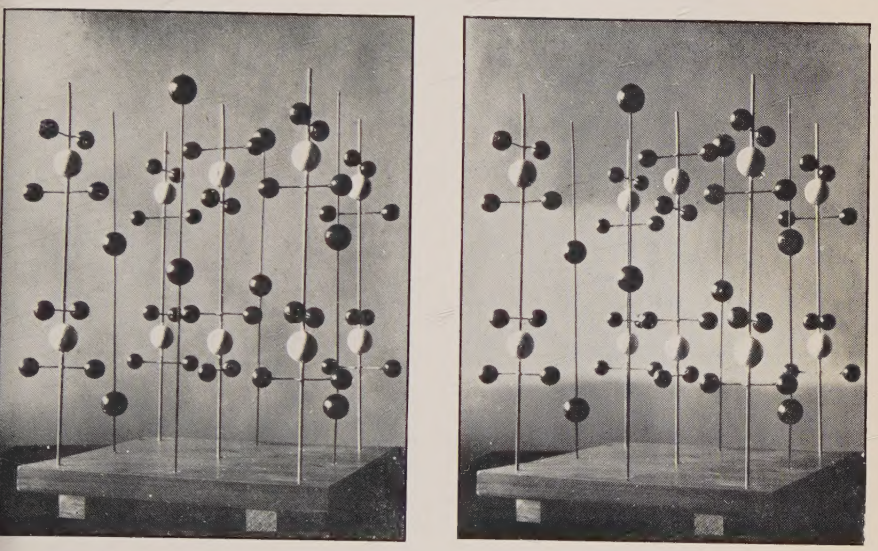


Fig. 7.